Essays on Social Networks in Development

by

Chutima Tontarawongsa

Department of Economics
Duke University

Date: ____________________

Approved:

____
Erica Field, Supervisor

____
Attila Ambrus

____
Rachel Kranton

____
Duncan Thomas

Dissertation submitted in partial fulfillment of the requirements for the degree of
Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2014

Abstract

Essays on Social Networks in Development

by

Chutima Tontarawongsa

Department of Economics
Duke University

Date: ______________________
Approved:

____
Erica Field, Supervisor

____
Attila Ambrus

____
Rachel Kranton

____
Duncan Thomas

An abstract of a dissertation submitted in partial fulfillment of the requirements for
the degree of Doctor of Philosophy in the Department of Economics
in the Graduate School of Duke University
2014

Abstract

This thesis aims to contribute to the understanding of the role of social networks in the context of developing countries. It contains two chapters that take on different aspects of social networks.

In the first chapter, I study a characteristic of social networks, network structure, which is the way in which households are connected in a social network. This chapter looks at how social networks help facilitate cooperation among people in the community, at least, in the context of contributing to public goods. Public goods in this case refer to shared infrastructure in rural villages such as community buildings, water wells, and roads. Provision and maintenance of these public goods relies on contribution of people within the village, and therefore is essentially a problem of voluntary collective action. I study whether the level of network connectedness of a household (as measured by network centrality) affects its decisions to contribute to public goods, using data from the Gambia. For an identification strategy, I use an instrumental variable approach that exploits the arguably exogenous variation in village ethnic composition, largely determined by historical accident. The findings suggest that better-connected households contribute more to some public goods. The network position effect is smaller when using a centrality measure that accounts for indirectly connected households. This paper also offers networks as a potential mechanism that explains the long-established relationship between ethnic diversity and public goods.

The second chapter (co-authored with Alessandro Tarozzi and Aprajit Mahajan) looks on a different aspect of social networks. In this chapter, we describe evidence of limited diffusion of bednet acquisition and usage from beneficiaries of an ITN distribution program in rural Orissa, India, to households that did not receive bednets during the intervention. Identification of such network effects relies on the change in ITN adoption among the beneficiaries of a program of bednet distribution that was carried out in a randomly selected subset of 141 study villages. This field experiment was designed to increase the adoption rate of insecticide-treated bed-nets to protect against malaria. The program randomly assigned 141 sample villages into 3 experimental arms a group in which some households received free distribution of bed-nets, a group in which micro loans for bed-nets were made available, and finally a control group with no intervention. In this paper, we focus on the impact of the intervention on households who lived in these respective groups of villages but did not receive the intervention. Our sample households include those that were exposed to the program via interactions with treated households. Identification is possible by exploiting the exogenous variation from the randomized controlled trials. We find that there is a small positive association between the number of social connections with treated households and their bed nets usage. On average, spillovers were limited. However, we find that bednet usage (but not acquisition) was substantively and significantly associated with some (but not all) measures of social links between non-beneficiaries and beneficiaries. This provides evidence, although limited, of network effects in the adoption of a health-related technology possibly due to diffusion of information and peer imitation.

For my first teachers, Mom and Dad

Contents

List of Tables

List of Figures

Acknowledgements

It would not have been possible to write this thesis without the help and support of my supervisors, Erica Field, Duncan Thomas, Attila Ambrus, and Rachel Kranton. Special thanks to my advisor, Erica Field, for her advice and for giving me the opportunities to do fieldwork in India and the Gambia. These have been eye-opening experiences for me. I appreciate all her contributions of time, ideas, and support to make my graduate school experience productive and stimulating. Her passion for development research motivates me during tough times in the PhD pursuit. I am also grateful for Alessandro Tarozzi for guidance in early grad school years. His words of encouragement have kept me persevere.

I am indebted to Dany Jaimovich of Goethe University Frankfurt for his generosity to share with me his data and for help with my work in the Gambia. Without his help, this project would not have been completed. I thank Raymond Jatta who has taught me much about the Gambia and its people during my field visit. I am also grateful for Mr. Nyakassi Sanyang and Ms. Lolley Jallow of the Gambia Bureau of Statistics for hours of their time spent searching for old documents for me. I thank participants of the 2014 EconCon at Columbia, the 2011 and 2013 NEUDC, a departmental seminar at Goethe University Frankfurt, a departmental seminar at Lafayette College, and Duke IPHD seminars.

I would also like to thank the Duke Department of Economics for providing me financial support and a great environment to work and be in. I am grateful to be a part of the community and to have met my classmates who have gone through this long journey together. The Development lunchgroup has contributed immensely to my learning at Duke. It has been and will be a place for people who share interests in international development to discuss and formulate research ideas. The development reading group has provided a safe space for me to freely ask any questions. I very much appreciate the group's enthusiasm and intellectual curiosity. I would also like to acknowledge the hospitality (and free candies) from the staff of Social Sciences 213 and the EcoTeach.

I couldn't have survived graduate school without my friends, Lala Ma and Lily Liu. Thanks for always being my keepers. We kept each other sane through out this tough process. I will always treasure countless hours we have spent together. Thanks my roommate of these past five years, P'Poy, for taking care of me and being my big sister (for washing dishes too).

A heartfelt thank goes out to my boyfriend, Pat, for your unyielding support in all my endeavors, for being the source of hope and joy, and for understanding and loving everything that I am.

And finally, no words can express how lucky I feel to be a part of my family - my parents and my little brother as well as my extended family members. Thanks for being by my side throughout every step of my life and for making me realize that even if all else fail, they will always be there for me.

For any errors or inadequacies that may remain in this work, the responsibility is entirely my own.

1

Network Structures and Public Goods Provision: Evidence from Village Kinship Networks in the Gambia

1.1 Introduction

The economic theory of public goods generally predicts that, when public good provision is private, the free-rider problem arises.[1] Yet, we observe abundant informal cooperation for voluntary public good provision, especially in the developing country setting. One theory suggests that social capital may help mitigate this market failure through networks such as kinship connections.[2] Additionally, one strand of game theory literature has investigated how network structures and network positions influence public good contribution (see, for example, Ballester et al., 2006 and Bramoullé and Kranton, 2007). Despite the rich theory literature, there is a lack of causal empirical evidence that establishes the relationship between networks and public goods. The current study aims to fill this gap in the literature by investigat-

[1] I consider public goods that are relatively non-excludable such as roads and water wells that are shared by everyone in a given village. Individuals have incentives to free-ride by under-reporting their valuations and still gain benefits of public goods due to non-excludability.

[2] For a comprehensive survey of the social capital literature, see Durlauf and Fafchamps (2005).

ing whether networks actually play a role in facilitating cooperation and whether the network position of an agent affects his contribution level.

In this paper, I provide rigorous empirical evidence that network position, as measured by network centrality, does indeed affect household voluntary public good contribution. I show that better-connected agents contribute more, using data from rural villages in the Gambia. The empirical challenge of this estimation lies in the endogeneity of the network formation process. An observed relationship between network position and public good contribution could be spurious, because network position is correlated with some other unobserved attributes. To establish this causal relationship, I employ an instrumental variable (IV) method that makes use of arguably exogenous variation in village ethnic composition to predict household network position. More specifically, I instrument household network centrality with the fraction of households in the village that belong to the same ethnic group as a given household. In an ethnically diverse country, such as the Gambia, having more households of the same ethnic group in the village serves as a proxy for having more kinship connections, mainly because of the patrilineal culture. It is worth emphasizing that this variation is at the village-ethnicity level rather than the village-level as in a more commonly used measure of ethnic fractionalization.

The validity of this IV hinges on the patterns of settlements being exogenous to factors influencing public good contribution. For a given ethnic group, the fact that group size is larger in some parts of the country than in others is, by a large extent, determined by historical accidents. To address the main concern for household sorting into villages, I show that, within the same ethnic group, households with small and large ethnic representation are similar across all observed socioeconomic and demographic characteristics. Moreover, data from the Gambia census reveal that village ethnic composition has been historically determined and remained stable across time. In light of the fact that there is no sorting based on a wide range of observed variables, any remaining concern for the unobserved variables correlated with public goods is likely to be minor in this case. Additionally, I employ an extensive set of control variables to reduce possible omitted variable bias. Given the current set of evidence, the exogeneity assumption seems plausible at least in this context. In terms of the exclusion restriction, the anthropological and historical literature documents that ethnic identity in the Gambia is not as salient in everyday life decisions compared to other West African countries (no incidence of ethnic conflict, no language barrier, and so on); therefore, ethnic composition should not influence public good decisions directly.

To estimate the effect of network position on public good contribution, I use a unique dataset from rural Gambia that contains information on complete village kinship networks. These traditional rural societies have poor access to electricity and reliable water sources. They rely heavily on community involvement to provide and maintain public goods with some assistance from the governmental and non-governmental organizations. Public goods considered in this paper include roads, communal buildings, water wells, public education and public health; all of which are shared at the village level.

My main results suggest that central households voluntarily contribute more. When using degree centrality as a measure of household network position, I find that the effect is slightly larger than that of eigenvector centrality, possibly indicating that direct links are more important than indirect links in public good decisions (see the appendix for formal definitions of network terminology). The estimates are robust to controlling for various types of relationships outside the village and using different subsamples that exclude extreme ethnic majorities and minorities. At the village level, there is some evidence that denser networks sustain a higher level of contribution, but the relationship is inconclusive. I interpret the results as evidence that social capital in networks can be used to facilitate cooperation. There are several mechanisms that can explain this network centrality effect on cooperation such as monitoring, social pressure, pure altruism, and information diffusion. However, because of data limitations, I estimate the net effects of these mechanisms and do not distinguish among them.

To my knowledge, this paper is the first to provide field-based evidence outside of a laboratory setting.[3] It also offers an alternative interpretation to the long established results that ethnic fragmentation lowers public goods, thereby adding to the public good literature. The results provide suggestive evidence of social networks as a mechanism causing this relationship. In this paper, the data allows me to break down the analysis to the household level; this is useful because agents, even within the same group, may face different levels of incentives to provide public goods. This breakdown helps supplement the results in the literature that mostly use aggregate public good data.

This paper contributes to the growing empirical literature that estimates the effects of social networks on economic outcomes, especially in the developing country context. For example, many network studies have been particularly interested in peer effects;[4] however, only a few recent papers have taken network structure into account when considering public goods contribution.[5] In this work, I attempt to estimate the causal effects of network positions on an outcome, a relationship that has not been studied empirically using a non-experimental approach.

While my empirical strategy is specific to the unique context of the Gambia,

[3] The findings resonate with the results of an experimental study by Carpenter et al. (2010) who conduct a lab experiment to compare public good contribution under different network structures.

[4] An agent behavior is influenced by his/her peers' behavior. See, for examples, Conley and Udry (2010), De Giorgi et al. (2009).

[5] Jackson et al. (2012) examine favor exchange in networks and propose that a threat of losing common links can be used to support a seemingly fragile one-shot game. They provide empirical evidence that such structure of having multiple common links is prevalent in Indian village networks. Alatas et al. (2012) and Banerjee et al. (2013) take a structural approach to model diffusion of information also in the developing country context. Both find that central agents are better informed and better able to contribute to the diffusion of information.

the findings yield broader policy implications for development projects that require collective action within a community. Examples of such projects are the growing Community-Driven Development (CDD) programs led by the World Bank. (Wong, 2012)

The rest of the paper is organized as follows. Section 2 outlines potential mechanisms and discusses theoretical works that establish links between network structures and public good contribution. Section 3 describes data used in the paper. Section 4 explains the identification strategy and establishes the validity of the method. Section 5 presents main findings and robustness checks. Finally, Section 6 concludes.

1.2 Theoretical Motivation

The motivation of this paper is largely based on theoretical works that have established a connection between social networks and public goods. In this section, I discuss possible mechanisms through which network positions may affect public good contribution. I also review the game-theory literature that considers cooperative games in the network setting. Although most existing models do not apply directly to the current context, I discuss how some aspects might explain what can be observed in the data. In terms of theoretical predictions, there is no consensus in the literature on the relationship between individual network position and the equilibrium level of contribution. Theoretical results generally depend on assumptions made regarding the excludability of public goods and the nature of network links. An empirical analysis in this paper therefore aims to provide some guidance for future theoretical works.

In rural areas of developing countries, public goods often refer to infrastructure such as roads, schools, and water wells that are sometimes non-excludable in the community.[6] Agents, or in this case, households, have access to public goods regard-

[6] This is suggested in the focus group interviews discussed in Section 3.

less of their contribution and network connections.[7] In terms of network interactions, there seems to be some strategic complementarities between the contributions of two connected agents, since we observe a positive correlation between them. However, the correlation, by itself, might reflect correlated unobserved heterogeniety because of the underlying network formation process. The interpretation of such complementarity is further discussed below. In the absence of networks, the situation reverts back to the classical public good model that results in free-riding of agents.

Among others, Belhaj et al. (2012) develop a model that can be applicable to the current setting, in that they study a class of cooperative games in networks with linear best-reply functions.[8] In this model, contributions between connected agents are assumed to be strategic complements. Their theoretical results show that when network effects are small, an agent's equilibrium contribution is proportional to his/her centrality. This simple prediction is appealing for an empirical test; however, the proportionality depends on the size of a theoretical parameter and the overall shape of network structures. When the network effect is large, this relationship breaks down. The equilibrium solution may not exist without further assumption such as imposing an upper bound on possible contribution. Moreover, if the interactions between actions of connected agents are strategic substitutes, instead of complements, the equilibrium solution may involve complex multiple equilibria (Bramoullé et al., 2013). Therefore, testing these models directly requires collecting more data as well as assumptions that may not be reasonable nor testable.

Instead of viewing network relationships as strategic interactions, a different

[7] A branch of theory literature focus on *local public goods* in which goods are non-excludable only along a link but not to others. Working examples of such public goods are innovations and information. The interactions between the contributions of two connected agents are often assumed to be strategic substitutes. Relevant papers include Bramoullé and Kranton (2007), Bloch and Zenginobuz (2007) and Galeotti and Goyal (2010).

[8] An example of payoff function that might fit the current setting of global public good is a modification of Equation (1) in Ballester et al. (2006) with an additional term to capture non-excludability at the network level *i.e.* agents' payoffs also depend on the total contribution.

model by Wolitzky (2013) considers network connections as monitoring channels (agents can observe only behaviors of those connected to them). Under this setup, a group can sustain the maximum level of cooperation (contribution) using a grim trigger strategy.[9] The model also makes a similar prediction that central agents have a higher level of cooperation. The equilibrium characterization relies on the fact that an agent deviates from his maximum cooperation once he receives some information of any deviation from his connected peers. The information about deviation then diffuses along network links to other agents. Eventually, the entire cooperation collapses after the deviation of at least one agent. In rural villages, however, public good projects are generally village-wide activities. It is possible that a villager can observe those without links present at an ongoing project regardless of network structures. Therefore, kinship networks used in this paper may not be the best proxy for a monitoring network for public goods.

As presented above, there is no consensus in the theory literature on how network positions or, more specifically, network centrality, affect public good contribution. Instead of directly testing any specific model, which would require making untestable assumptions given the current dataset, I estimate the empirical relationship between network centrality and household public good contribution. This exercise can help shed light on the direction of the relationship and, indeed, whether such relationship exists at all.

The theoretical basis in the above set of literature motivates me to choose eigenvector centrality among many network measures in the literature.[10] While a simpler measure, degree centrality, focuses only on the number of direct peers; eigenvector centrality is designed to capture impacts from both directly and indirectly connected

[9] Each agent plays the maximum level of cooperation, until he observes a deviation.

[10] Eigenvector centrality corresponds to a special case of Bonacich centrality, the measure used in Ballester et al. (2006) and Belhaj et al. (2012). See the appendix for further discussion.

agents. It assigns centrality scores so that an agent's score is a weighted sum of the scores of his/her connected agents. Having a more central neighbor contributes more to an agent's score than having a less central neighbor. Another benefit with using the eigenvector centrality is that the results can be compared with other empirical studies that also use the same measure (for example, Banerjee et al., 2013, Alatas et al., 2012).

In addition to the theory literature, a similar question has been studied only in an experimental setting by Carpenter et al. (2010). In the experiment, subjects were assigned into groups of four and asked to play repeated public good games. In this context, the (exogenously imposed) network connections dictated who observed whom in terms of contribution. The researchers found that agents contributed more when being monitored more. Also, well-connected networks yielded a higher level of contribution.

In a non-experimental setting, there are other possible mechanisms that can explain why network centrality might affect public good contribution. I outline three possible mechanisms in addition to monitoring in this section.

First, central agents may face a higher threat of social punishment. They have more to lose if they shirk. The idea of social punishment or social sanction has been extensively explored in the public good literature; yet, it largely ignores social networks as an implementing channel. For example, Miguel and Gugerty (2005) build a simple public good model in which agents can punish free-riders of the same ethnic group but not across ethnic groups. Therefore, a highly diverse community has a lower contribution level because its members cannot impose punishment effectively. However, such a relationship cannot explain a low level of public goods in places with ethnic homogeneity (Esteban et al., 2012). On the contrary, these cases may be better explained by network fragmentation, rather than ethnic fragmentation. Consider two unconnected agents from the same ethnic group; if one free-rides, another would not have any credible means to assert social pressure. On the flip side, it is also possible that two agents of different ethnicities may be connected and thus possess the means to punish each other. This type of punishment is also observed in experimental settings (see Fehr and Gaechter, 2000, Masclet et al., 2003, Ertana et al., 2009). In some cases, findings indicate that experimental subjects choose to punish free-riders even at their own costs.

Second, network centrality may serve as a proxy for the true valuation of public goods if an individual valuation of public goods depends upon the number of friends/families he has. A central agent has more to gain by contributing to public goods because she cares about more people within a given network. This mechanism is closely related to altruistic links among extended family studied in the previous literature (Altonji et al., 1992, Cox and Fafchamps, 2008).

Third, the free-rider problem could stem from information asymmetry.[11] When the true valuation of public goods is private information, agents have incentives to underreport their preferences, resulting in inefficient allocations. In the network context, it is possible that the private information of central agents is better known within their networks. There is some evidence in models of information diffusion that more central agents are better informed (Banerjee et al., 2013, Alatas et al., 2012). The opposite may be true that their private information might also spread faster in the network.

Despite the rich theoretical literature, empirical studies that provide causal evidence have been limited. The current study provides empirical evidence for the relationship between network centrality and voluntary public good contribution. Due to the limitations of the data, I estimate the net effects of the mechanisms outlined above and do not distinguish between them.

[11] This is studied in the literature related to Vickrey-Clarke-Groves mechanisms.

1.3 Data

1.3.1 Survey and Sampling method

I examine the effect of network positions on voluntary public good contribution using cross-sectional data from the Gambia. The data is part of a baseline household survey for an impact evaluation of the Gambia Community-Driven Development Project (CDDP) conducted between February and May 2009.[12] To create a nationally representative dataset, wards were randomly selected from 6 Local Government Areas (LGA) out of 8 LGAs covering the entire country. Then, 3 - 8 rural villages were randomly chosen from each selected ward. Table 1.1 shows summary statistics at the village level. The villages are small with the number of households, ranging from 18 to 138. The economic conditions are those of traditional rural societies with little access to electricity and improved water sources.

The data contain information on 2,718 households from 60 villages in rural areas.[13] The aim of the survey was to cover a wide range of networks as extensively as possible. This includes data on complete networks for various types of relationships within the village. Using an ethnographical approach, the survey team covered 94% of all households in the sampled villages.[14]

For the main analysis, I restrict the sample to a subset of villages referred to as *stable villages* because my identifying assumption is more credible for this given subset. The subsample includes 2,110 households in 48 villages.[15] The details regarding

[12] The data were generously shared with me by Dany Jaimovich of Goethe University Frankfurt, Germany. Jaimovich (2011) and Jaimovich (2013) extensively analyze this dataset focusing on network formation and substitutability between links within and across villages.

[13] The network module in the survey is most complete with information of networks on 2,886 households as in Jaimovich (2011). The final sample is smaller due to some missing variables. Even though I need to drop some households with missing data, they still count as a part of networks of households in the final sample

[14] This approach requires gathering household heads together to collect information. See Arcand et al. (2010) for more detail on data description.

[15] A *stable village* refers to a village in which ethnic group composition has been stable since 1993

this restriction will be discussed in Section 4. The results for the full sample are similar to the main analysis and are included as a robustness check.

In rural Gambia, villagers are usually organized into compounds in which members of the same family, related by blood or marriage, live together in a group of huts. Usually a compound can be identified as a household. In some cases, however, some members of a compound declare themselves as an independent household (15% of the sample household heads are not compound heads). The distinction between a household and a compound is generally recognizable to the village head and other villagers.

Table 1.2, Panel A presents descriptive statistics for households in the sample. The large average household size of 13 is explained by polygamous culture, with 48% of the households being polygamous. A head is generally the oldest male in the household or his widow. A typical household may consist of multi-generational nuclear families (van de Walle and Gaye, 2006). The average annual self-reported per capita income was 2,282 Gambia Dalasi (approximately 302 PPP-adjusted US Dollars per year or 0.8 US Dollars per day in 2008). This monetary income was likely an inaccurate measure of standard of living of these households because of agricultural activities and the informal sector.[16] The preferred measures of wealth are the number of huts with corrugated roofs and the number of huts with grass roofs.

1.3.2 Public goods in rural Gambia

The setting of a traditional rural society makes the Gambia suitable for my interest in network-level interactions. The cooperation of villagers in providing and main-

(the earliest census data available).

[16] Since the majority of households (73%) worked in the agricultural sector, economic transactions could be in-kind. Their actual consumption and standard of living may have been higher than the self-reported monetary income. Unfortunately, we do not have information on consumption or sources of income.

taining public goods is prevalent, although there is some involvement of government and NGOs. This dataset contains unique information on individual household contribution that allows me to study household-level incentives to contribute to public goods.

The questionnaire asked whether each household had any member who participated in specific *village activities for the community* including roads, communal buildings, water wells, public education, and public health. Even though the question does not ask for specific tasks performed as a contribution, public goods contribution in this case is defined broadly to include maintenance of the goods as well. In addition, the survey specifies the form of each contribution (e.g. voluntary, paid work, or monetary contribution). Approximately, 6% of all contributions were monetary, less than 1% was work with pay, and the rest were voluntary work. In this study, I focus on voluntary contribution including both monetary and in-kind work and exclude paid work to isolate the incentives to contribute induced by networks rather than monetary incentives.

Ideally, we would measure contribution in a more detailed manner, *i.e.* total monetary amount or time spent contributing. However, the survey elicits only binary variables for contribution to different categories of public goods. In addition to these binary variables, I construct another outcome variable by summing the total number of public goods to which a household contributes. Panel B of Table 1.2 shows summary statistics of the six outcome variables. Note that only 3% of the sample households reported not contributing at all. On average a household contributes to 3.1 public goods.

The pattern of contribution reveals that there is a high correlation for contributions among households in the same village. For each type of public goods, there are some villages with zero contribution. Since the timing of the public good question covered only one year prior to the survey, this pattern might indicate that contribution also depends on projects occurring during the time period. This circumstance, in turn, depends on needs that arise in the village. Since, the definition of public good contribution is defined broadly, I assume that there is always an opportunity for villagers to contribute to any type of public goods.[17]

In addition to the main questionnaire, the survey team conducted focus group interviews among villagers in 29 villages (as reported in Arcand et al., 2010). The participants were selected at random and invited to the interviews.[18] Even though we do not have specific information about existing public goods or the need for certain public goods, this focus group report provides some insight into the public good situation in the rural Gambia context. The report suggests that, in all 29 cases, both male and female villagers (who composed separate interview groups) considered access and distribution of public goods to be universal. Given this response, public goods considered are likely to be non-excludable at the village level.

I next summarize some relevant aspects of each village public good as described in the field report. First of all, water wells were noted as most inadequate in 27 out of 29 villages, followed by roads, public health, and public education. Villagers typically shared one or two wells/boreholes with a hand pump. Based on the government report on water resources (Jarju, 2009), these wells were hand-dug and not very deep because of the shallow sand aquifer that received underground water from the Gambia River. Given that every household needs water and that water wells are typically the main source of clean water, they are a good representative of public goods that require voluntary contribution. The second most important public good is the availability of roads. Having good roads is related to having improved access

[17] I also experiment with conditioning on whether a public good project exists by excluding villages with zero contribution for each type of public goods. Doing so does not significantly change the main results.

[18] In practice, this was not always the case because some villagers passing by might have joined in as well.

to hospitals, markets, and other facilities outside of the village. This access is crucial to monetary income because the rural Gambia economy relies heavily on groundnut exports (Sylla, 2010). In the Gambia, there are two main all-weather roads that stretch from East to West along the north and the south banks of the Gambia river. Smaller roads that branch out into villages are generally dirt roads that can be easily damaged during the monsoon season. The baseline report also suggests that some villages were not accessible by roads all year round. Contributions to roads may include activities such as building a new dirt road and repairing damages from flood and rain. Another type of public good is communal buildings, such as town halls. A mosque is also considered a community building; therefore, incentives to contribute might vary with religion. This motivates me to control for religion in all specifications. Note that 99% of the sample population is Muslim.

Public education and public health refer to the construction and maintenance of related facilities such as schools and village health facilities. A specific example, as told by a village chief during an interview: an NGO recently started a primary school project by supplying construction materials, the villagers then had to cooperate and organize their own manual labor for the construction. There could also be other types of contributing tasks that require certain skills such as teaching or medical training. More educated individuals may be more able to contribute to public education. For this particular example, however, since the majority of the sample is illiterate, contribution to public education may not necessarily mean formal teaching. For instance, Gamble (1949) documents the use of folklore as a way to transfer knowledge. Given this type of education, one does not require literacy to be able to contribute to public education.

1.3.3 Network Data

The network module contains binary data on kinship ties among all households within a village. All relationships considered here are undirected and unweighted.[19] The type of relationships considered here is kin networks, a term that refer to the blood relations of a household head and his wives.[20] The questionnaire asked which households the household head and wives had direct blood relationships with (father/children, direct brother/sisters).

Figure 1.1 illustrates network graphs that represent kin relationships among households in 2 villages in the sample. Each node represents a household and each line presents a kin relationship between the two connected households. The colors of the nodes show whether the household contributes to a particular type of public good (as labeled in the figure).

Two different measures of network centrality are selected based on the theoretical motivation as discussed in Section 2. Each of which measure intends to capture the relative importance of a node in a graph under different concepts. The simplest measure of centrality, "degree centrality," is the number of links a household has. This measure, however, does not take into account other aspects of the network beyond the locality of a given agent. It is possible that being connected to an influential agent may make one more influential relative to being connected to a non-influential agent. "Eigenvector centrality" captures this aspect by weighting in the centrality of each connected node. This measure allows households with the same degree centrality to vary depending on their extended networks. The correlation

[19] An *undirected* link means that household i and household j are connected even if only one of them reports the relationship. An *unweighted* link refers to a binary relationship with no information on the strength of the tie.

[20] The data also contains information on networks of economic-interactions among households. I focus mainly on the kinship networks for which my identification strategy is strongest. The results for economic-exchange networks are also included in the Online Appendix.

between the two measures is approximately 0.6.[21] Eigenvector centrality differs more from degree centrality when there are many low-degree nodes (households with fewer links) connecting to high-degree nodes. Figure 1.2 shows the distribution of degree centrality and eigenvector centrality.

Additionally, at the village level, I present a descriptive analysis of how aggregate public good contribution (the percentage of households contributing to public goods) may vary with a network structure measure. I use network density, which represents how well connected households are within a village relative to all possible connections. The descriptive statistics for all network characteristics are reported in Table 1.3.

1.4 Estimation Strategy

1.4.1 Identification

In estimating the effects of network position on public good contribution, the main concern is the endogeneity of network positions.[22] Households choose with whom to form links. The direction of the bias in OLS estimates is ambiguous depending on the underlying network formation process. Furthermore, there may also be other omitted variables. For example, a previous study by Anderson et al. (2004) identifies trust as a determinant reducing individual public good contribution. If more trustworthy households are also more central, omitting a trust measure may therefore bias the OLS estimates downward. Another determinant of public good contribution that may also create a downward bias is the number of females in the households. Nowell and Tinkler (1994) find that females generally contribute more to public goods. In a male-dominant society, having more females in a household might be associated with being less central in kinship networks. On the other hand, the direction could

[21] This is significantly lower than the correlation found in Valente et al. (2008) who consider 62 health-related networks from 7 different studies.

[22] See Manski (1993a) for a discussion of endogenous network formation.

be the reverse, if we consider a potential reverse causality that contributing more would attract more connections. Since none of these factors is particularly dominant in this context and the network formation process is unobserved, the direction of bias is left as an empirical question.

To address the endogeneity of network centrality, I exploit *household ethnic fraction* within a village as an instrument for household network centrality. This variable is defined as the fraction of households within the village that belong to the same ethnic group as the current household. The ethnicity of a household is the ethnicity of the household head which is pre-determined by that of his father. For illustration purposes, in a village of 10 households, if 3 households belong to the same ethnic group, their associated value of the *household ethnic fraction* is 0.3.

Since a kin link is more likely to form between households of the same ethnic group due to the patrilineal culture, *household ethnic fraction* should be a strong predictor of household network centrality. Note that since kin networks are defined as the blood relatives of both husband and wives, it is not necessary that two connected households must be of the same ethnic group. In fact, about 17% of all kin links are between households from different ethnic groups due to inter-ethnic marriages. I use the fact that rural Gambian villages are highly ethnically diverse, and that this variation is arguably exogenous to estimate the effects of household network position on public good contribution.

The validity of this IV method rests on the village-level patterns of settlements being exogenous to the factors that influence household public good contribution decisions. The identification is derived from comparing households of the same ethnicity who reside in places with variation in *ethnic fraction.* The main concern is household sorting that leads to selection of different household types into villages where they have less or more ethnic representation. I provide evidence against this concern by showing that, within ethnicity, household observable characteristics are uncorrelated with household ethnic fraction. The key identifying assumption is that, for a given ethnic group, group size (as measured by ethnic fraction) is bigger in some parts of the country than in others, and this variation is, to a large extent, determined by historical accident. Evidence supporting this exogeneity is discussed extensively in the next subsection.

The exogenous treatment of ethnic composition has been used in the previous literature. Prominent papers, including Easterly and Levine (1997) and Acemoglu et al. (2001), treat ethnic fragmentation at the cross-country level as exogenous when studying it as a determinant of economic growth. Among those who attempt to tackle the endogeneity of local ethnic composition, Miguel and Gugerty (2005) use residential ethnic composition to instrument for school-level composition in Kenya. In the current context, I make an argument similar to that in the Kenya case: that residential ethnic composition at the village-level has been historically determined and that contemporaneous migration is limited.

Figure 1.3 displays a reduced-form relationship between *household ethnic fraction* and household total public good contributions.[23] I hypothesize that this positive relationship develops through the effect of household ethnic fraction on household social networks and subsequently through social networks on public good contribution. To allow for non-linearity, I also present a kernel-weighted local polynomial smoothing estimate in the same figure. Overall, the nonparametric estimate is similar to the linear estimate, although the ethnic fraction effect increases slightly between 0.4 and 0.5 but tapers off at 0.7.

In the rest of this section, I provide some background detail on ethnic groups in the Gambia. Then, I provide evidence to support the exogeneity of the instrument. Finally, for the exclusion restriction, I discuss the possibility of mechanisms other than networks through which household ethnic fraction may affect public good

[23] Full reduced-form results with all 6 outcome variables are included in the Online Appendix.

contribution.

1.4.2 Validity of the IV

Despite the country's small size, Gambian population comprises many diverse ethnic groups. Major ethnic groups include Mandinka, Fula, Wollof, Jola, Serahuleh, Serer, and Manjago. In this data, the average village ethnic diversity index (Herfindahl) is 0.3.[24] Historically, these ethnic groups migrated to Gambia from various parts of West Africa, with Serer being the oldest inhabitants in the area. The dominant group is Mandinka, who migrated to the Gambia during the spread of the Mali Empire in the 13th century. The Wolof is generally located in the Western part of the Gambia River. An extensive study series by anthropologist David Gamble (1949) explicitly states how the Wolof occupy roughly the same area now that it did in the 15th century. Even the newest ethnic group, Serahuli, arrived from the larger Senegambia area since the 1800's (Mwakikagile, 2010). Across all villages in the sample, those that are closer to the capital, Banjul, generally have higher ethnic diversity, possibly due to external migration to the economically developed urban areas. Also, among different the ethnic groups, six different ethnicities form the village majority group at least once in this data.

The census data, collected by the Gambia Bureau of Statistics (GBoS), reveals that ethnic composition at the national level has barely changed since 1983, indicating that settlement patterns have been historically stable or that the change has been extremely gradual. Between 1983 and 2003, the largest group, Mandinka, remained at approximately 34-36 % of the total population; the second largest group, Fula, made up 17% with an increase to 22% in 2003.[25] Note that the dataset oversampled Mandinka and undersampled Wolof and Jola.

[24] According to Fearon (2003), Gambia's ethnic fractionalization index is 0.76 and is ranked the 26th most diverse country in the world.

[25] More detailed data are provided in the Appendix.

In order for household ethnic fraction to be a valid instrument, I argue that households from the same ethnic group settle in places with variation in their ethnic fraction by random historical accident. This claim is supported by the data. The main concern is that there may be a selection process that leads to different types of households sorting into villages related to public goods. In the context of rural Gambia, sorting in this way seems to be limited. To investigate whether this type of sorting is relevant in this setting, I compare households from the same ethnic group with different ethnic fractions across all observed characteristics. The balance check in Table 1.4 shows that they are similar across the categories of total income, income from agricultural activities, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of household workers, size of land owned, whether land is of high quality, role as the village chief, role as a traditional healer, role as a village development committee member, role as a marabout, religion, marital status, relative wealth, number of corrugated huts and number of grass-roof huts. Of all variables, only one (having an imam in the household) is significantly correlated with household ethnic fraction.[26] Such an occurence is likely to be random. The results reinforce the argument that ethnic fraction of households within the same ethnic group is random and likely to be determined by historical accident.

In all specifications in this paper, I also control for this extensive list of variables to correct for any potential imbalance. The remaining concern that may invalidate the instrument is household sorting based on unobserved characteristics correlated with pubic good contribution. Based on the current results and further information from the census data, this concern seems to be minor in this case.

One implicit assumption that I make in the analysis is that only networks of

[26] Excluding households with imams (approximately 20 households) does not change the main results.

households within the same village matter for public good decisions. For kinship networks, the ethnic fraction variable serves as a proxy for extended families residing in the same village. Households with a smaller ethnic fraction could have a large kinship network living outside of the village; however, since public goods are shared at the village level, networks outside the village should not affect their within-village public good decisions.

A further investigation using the 1993 and 2003 census data shows that the ethnic fraction variable has been stable in the sample villages over the ten-year span. Recall that the full sample contains data from 60 villages. I could confirm the stability of this variable in only 48 villages and therefore restrict my main analysis sample as such. At the village level, the absolute percentage change of each ethnic group has an average close to zero for all groups. 8 out of 60 villages have one or two groups with a relatively large percent change.[27] Most of these unstable villages are located closer to the Banjul areas, especially in the upper and lower Niumi districts across the river from the capital. In particular, one unstable village that lies on the West coast near Banjul experiences a drastic fall in the "Others" ethnic category. Another unstable village has a 275% increase in population size. This change is not true in the stable villages. As a robustness check, I perform an analysis on the full sample as well.

Note that the changes in the census data are at the population level rather than at the household level as used to construct the IV. To some extent, these changes could reflect household expansion rather than creation of new households (which would not affect the ethnic fraction variable). Unfortunately, the number of households in each village is reported at an aggregate level that combines several nearby villages into a 500-household "census settlement." Villages forming each settlement also vary between the two census rounds, making it difficult to determine whether the observed

[27] Data on 4 villages were not available in the 1993 census.

population changes are due to migration or household expansion.

Additional data from the 2003 and 2009 migration reports address the concern about migration in response to public goods. Mobility in rural Gambia is generally low. The 2003 Migration Analysis reports that 96% of internal migration in the Gambia is rural to urban migration. The common destinations are the urban Kanifing and Banjul areas, which are not included in the sample. The 2009 Migration and Urbanization survey suggests that the main reasons for migrating are marriages and work (36.4% and 19.8% respectively). The majority of migrants are either women who marry men who live the Banjul area or young men who move to find employment in the city. Given that an average household size is large, this type of migration may change household structure but has minimal impact on household ethnic fraction in rural villages. As discussed earlier, the household-level census data is not available to test this hypothesis; however, in interviews with 2 village chiefs (in 2 rural villages outside the sample), the chiefs did not recall any situation where a migration involved an entire household.

Additionally, ethnic composition may affect public good contribution via channels other than networks. It has been established that ethnic fragmentation lowers public goods in the previous literature, although only a few studies distinguish between mechanisms explaining such a relationship. Habyarimana et al. (2007) attempts to do so using a lab experiment in Uganda, finding find strong evidence in favor of mechanisms that go through networks. I next discuss and provide evidence to rule out channels other than networks in the current context.

First of all, one of the most studied mechanisms suggests that conflicts or discrimination between ethnic groups can be the main cause of lower public goods. It is well documented in the historical and anthropological literature that ethnic groups in the Gambia live together harmoniously relative to other West African countries such as Cote d'Ivoire, Ghana, or Mauritania (Sonko-Godwin, 2003, Mwakikagile, 2010). There is low to no tribal conflict. Despite different origins, the colonization of the British, the abolition of slave-trade, and intermarriages between different ethnic groups have promoted cultural integration among the ethnic groups. There is also evidence against economic tension between ethnic groups. Arcand and Jaimovich (2012) use the same data as does the current study to show that the difference in ethnicity does not hinder households from engaging in an economic interaction. The impact of ethnic diversity on public goods through direct conflict seems irrelevant in this case.

Another possible channel is that minority households may have less access to public goods and hence may not contribute. Even though I cannot directly test this in the data, according to the focus group interviews, all groups in the 29 interview villages report that access to public goods is universal. Moreover, even if the distribution of public goods is dictated by a village authority figure, such as the village chief and village development committee, being a village chief is not correlated with being from a large or small ethnic group, as evident in the balance check in Table 1.4.

Households of the same ethnic group may possess technology that enhances cooperation among them. An example of such technology is language. People of different tribes may have a language barrier in cooperating for public goods. However, this is not the case in the Gambia. In general, a large fraction of population - including children - speaks more than one language, especially Mandinka. More importantly, individuals can speak the language of the dominant tribe in their own villages. Formal schooling is conducted in English. Another type of technology that may enhance collective actions is social sanction (see, for example, Besley and Coate, 1995 and Miguel and Gugerty, 2005). This requires an assumption that households can only sanction those in the same ethnic group. It is more credible to specify a precise channel, such as network links, as a way households can punish one another.

Last but not least, another explanation suggests that households may simply have altruistic feelings strictly toward others in their own ethnic group. If true, we should see that, at the village level, higher ethnic fragmentation reduces the aggregate level of public goods because households feel less altruistic towards others in the village. Panel A of Table 1.5 shows that when the village-level contribution is regressed on ethnic fragmentation index (Herfindahl index), the coefficients are not significant for all types of public goods. The coefficients are small with large variances. The same results hold when controlling for network density in Panel B. Moreover, the coefficients for network density are much larger than the coefficients for ethnic fragmentation index. In addition to the village-level evidence, there are also other reasons to believe that this channel is a minor concern. In the context of the Gambia, ethnic identity is not salient in everyday life decisions as manifested in the lack of ethnic conflict incidence, the lack of language barrier, and so on. The spread of Islam throughout the region reconciles differences among people and seems to create an identity that is more unifying than ethnic identity (Mwakikagile, 2010).

Despite indirect evidence, altruism strictly toward co-ethnic households remains a potential limitation for the exclusion restriction of the IV. Overall, since channels other than networks do not seem relevant in this context, the exclusion restriction of the IV, household ethnic fraction, is plausible.

1.4.3 Empirical specification

For the household-level analysis, I estimate the effects of household network position on public good contribution using a 2SLS regression. The first-stage specification is as follows:

$$Centrality_{ijk} = \alpha_0 + \alpha_1 EthFraction_{ijk} + \alpha_2 X_{ijk} + ETH_j + v_k + u_{ijk}$$

where the IV, $EthFraction_{ijk}$, is the fraction of households in the village k that belongs to the same ethnic group j as household i, $Centrality_{ijk}$ refers to the two measures of household network centrality (degree centrality and eigenvector centrality), X_{ijk} is a vector of household characteristics, ETH_j are ethnicity fixed effects, v_k are village fixed effects, and finally u_{ijk} is the error term for the first stage regression.

Table 1.6 presents the first stage regression results. The coefficient, α_1, is highly significant. *Household ethnic fraction* works well in predicting kin network centrality.[28] The F-stats are 61.03 for degree centrality and 94.45 for eigenvector centrality, well above the Stock and Yogo (2005) critical value of 16.38.

The second stage is then:

$$Contribute_{ijk} = \beta_0 + \beta_1 \widehat{Centrality}_{ijk} + \beta_2 X_{ijk} + ETH_j + v_k + \epsilon_{ijk}$$

where $Contribute_{ijk}$ is the outcome variable (5 binary variables for different types of public goods and 1 non-binary variable, total contribution), $\widehat{Centrality}_{ijk}$ is the predicted network centrality from the first stage, and ϵ_{ijk} is the error term in the second stage regression.

The main coefficient of interest is β_1, which represents the effect of network centrality on household public good contribution. The outcome variables are binary (with an exception of total contribution), so β_1 can be interpreted as a probability.[29]

In all but the first set of the OLS estimates, I control for a range of household characteristics (see the list in Table 1.4). Because my analysis is at the household level, household size and the number of household workers might be problematic: having more members and workers both relaxes the household resource constraint to contribute to public goods and increases the channel of connections to other

[28] The first stage regressions are similar to a specification in Table 8 of Arcand and Jaimovich (2012).

[29] Probit model yields similar results in comparisons to the linear probability model. The probit results are available upon request.

households. Controlling for these variables is crucial. However, the effect of having an additional household member on network centrality may not be constant for all levels of household size. To identify networks separately from the ethnic composition of the village, I experiment with a more flexible form of the household size variable by using household size relative to the rest of the village (ranking) instead. I also control for village fixed effects, because we do not have information regarding either existing public goods in each village or their specific needs for public goods. This control also sweeps out common factors such as availability of projects and government/NGO involvements in providing public goods. Since the sample villages are relatively small, it is reasonable to assume that all households face the same public goods within the village.

1.5 Effects of Network Centrality on Public Good Contribution

1.5.1 Descriptive Results

I begin by presenting the descriptive results.[30] Following the theoretical motivation laid out above, I consider two different measures of centrality: degree centrality and eigenvector centrality. Degree centrality counts only the directly connected households, whereas, eigenvector centrality captures a broader network structure beyond the given household's locality. Table 1.7 shows how the two centrality measures vary with observed characteristics. Both measures are correlated with several variables, indicating that network formation is likely to be a non-random process.

Next, I analyze how network density is associated with aggregate public good contribution at the village level. Intuitively, a denser network has higher social capital that could better facilitate cooperation, leading in turn to higher levels of public good contribution. Some of the frameworks in the theory literature discussed in Section

[30] In all analyses, I drop households that report unusually large degree centrality. The cutoff is at 15 which is approximately the 99 percentile. There are only 20 households above this cutoff. Using the full sample does not change the results by much. The results are available upon request.

2 also extend to predict network-level contributions. The outcome variables are percentages of households in the village that contribute to different types of public goods.

Panel B of Table 1.5 shows that a denser network is associated with more contributions for roads, public health, public education, and average total contribution. The relationships are statistically significant for all of above but public education. The results are reversed for contribution to communal buildings and water wells. Surprisingly, for these two types of public goods, sparser networks are associated with more contributions. This set of results, however, cannot be interpreted causally, because the results may be driven by omitted variable bias. For example, a high network density may capture the fact that households live relatively closer together (despite the same level of village area) and therefore need fewer contributions to water wells. More data is needed to disentangle the effect of network density on aggregate contribution. Nonetheless, this descriptive result helps to bridge the main results at the household level to community-level public goods, for which data are generally more available in other contexts. The village-level evidence also allows for comparisons with the previous literature that studies public goods at the aggregate level (for example, Alesina et al., 1999, Miguel and Gugerty, 2005).

Before proceeding to the IV results, I establish correlations between network centrality and voluntary household contribution to public goods. Panel A of Table 1.8 reports the OLS estimates when regressing the six contribution variables on the two network centrality measures. Each outcome variable is regressed separately on each centrality measure. Overall, both centrality measures are positively correlated with public good contribution. More specifically, we see significant and positive correlations for contribution to community buildings, water wells, public education, and total contribution. It is worth emphasizing that the two centrality measures are of different scales.

Panel B of Table 1.8 reports similar specifications while holding a range of household characteristics constant. The coefficients do not change by much, even though the centrality measures are correlated with some of the control variables. The coefficients of both centrality measures for total contribution become slightly smaller and less precise. Some of the household characteristics are significantly correlated with contribution, but there is no obvious pattern across public goods.[31] Farmers are more likely to contribute to roads. More education is associated with contributing to community building construction. Households with fewer members and more workers are more likely to contribute to water wells. Households with traditional healers also contribute more to water wells as well as public health. Households with marabouts, on the contrary, contribute less to public health. Those in the village development committee contribute more to public education. Finally, the village development committee and farming households contribute to more types of public goods overall.

1.5.2 IV Results

In Table 1.9, the IV results suggest that network centrality increases household contribution to public goods. The coefficients are positive for all types of public goods, but they are significant only for water wells and total contribution. In all cases, the IV estimates are at least twice as large as the OLS estimates. The downward bias indicates that such factors as a reverse causality from public good contribution to centrality may not be as dominant.

The effect of degree centrality is particularly strong for water wells. The marginal effect of having an additional kin link increases the likelihood to contribute by 3%. The difference between an isolated household and a household with an average num-

[31] I only present the coefficients of network centrality but omit estimates for household characteristics.

ber of links, 4, is 11%; the size of the effect accounts for 25% of the mean water well contribution. The effect is as large as 75% for a household with the highest degree centrality (15 links). For total contribution, the network centrality effect is small compared with the mean outcome.

The eigenvector centrality measure diverges from degree centrality by taking into account indirectly connected households. Eigenvector centrality also allows households with the same degree centrality to vary depending on the extended network structure surrounding them. This measure can be interpreted as a weighted sum of *walks* originated from a given household to other indirectly connected households.[32] A longer walk (between the originated house and an indirectly connected household) is discounted more. Thus, a household is more eigenvector-central, if it is connected to more households with shorter distances.[33] The distribution of this measure, as shown in the Figure 1.2, has a fatter tail above the mean (0.79) compared to the distribution of degree centrality, with a large number of households clustered close to 0. The shape of the distribution indicates that when the global network structure is taken into account, the difference between the mean and the most central households is wider than is the difference when one focuses on the local network using degree centrality.

Since the value of eigenvector centrality is relative, it is more meaningful to interpret the magnitudes of the coefficients by direct comparisons. Increasing household eigenvector centrality from 0 to the 25^{th}, 50^{th}, 75^{th}, and 99^{th} percentiles raises the likelihood of contributing to water wells to 3%, 8%, 15%, and 41%, respectively. Notice that the increment is much larger at the higher percentile because of the distribution shape. For comparison purposes, the same calculation for the degree centrality estimates yields 6%, 11%, 17%, and 39%. Overall, the magnitudes of

[32] A *walk* is a sequence of links connecting a sequence of nodes (households).

[33] See the terminology appendix for further explanation.

the eigenvector centrality coefficients are slightly smaller than those of the degree centrality, possibly suggesting that direct links are more important to public good decisions.

Despite positive effects for water wells and total contribution, the lack of coefficient significance of other types of public goods does not imply that there is no relationship; it may be a result of the lack of statistical power. When the sample is expanded to the full sample, at least two other coefficients become significant (community buildings and public education). Yet, the centrality effects for roads and public health seem especially low. In this context, road construction and public health may rely more on public provision from the government than from household cooperation. The field report suggests that, from the perspective of villagers, roads as a public good generally refer to roads that connect one village to another, as opposed to those used within a village. In this case, contributing to roads may be less feasible for a private provision project to undertake and may require cooperation across villages. Similarly, public health may also be difficult for within-village provision, because providing public health may need constant outside support from professionals with formal healthcare training.

1.5.3 Robustness checks

The validity of the IV results in Table 1.9 depends on the exogeneity of the IV. As argued above, the assumption seems plausible, at least in the current setting. I further confirm the results by performing three robustness checks. Overall, I find that the main IV estimates are robust with small changes in magnitude.

First, I add another set of controls, including network relationships outside the village and enumerator fixed effects. In addition to the within-village networks, the dataset contains information on the existence of links with households outside the village. The information is, however, limited to binary variables - whether the household has different types of relationships outside of the village rather than the actual number of such external links. These variables were previously excluded for parsimony. The types of external links include having a credit exchange (borrowing or lending), a labor exchange, a land exchange, a productive input exchange, and lastly marriage relationships. Furthermore, I also attempt to control for measurement errors that may arise from data collection by controlling for enumerator fixed effects. During the survey, respondents were in interviewed in groups. One concern is that the respondents might be influenced by the enumerator and other respondents present during the same interview.

Panel A of Table 1.10 reports the IV results for this robustness check. Compared with the main results, the estimates do not change by much when more controls are added. Most coefficients increase slightly, but do not differ significantly from the previous estimates. Only in a few cases are the enumerator fixed effects significant, indicating possible but limited measurement errors from interviewing in groups.

As a second robustness check in Panel B, I use the full sample to ensure that the results are not driven by the selection of the stable village subsample. When the full sample is used, the results show a similar pattern to the main analysis with only stable villages. The magnitudes are slightly larger. The coefficients of network centrality become significant for communal buildings and public education.

Panel C of the same table presents the results for another subsample that excludes extreme cases of households with the largest and smallest ethnic fraction. These extreme ends of the ethnic fraction distribution may capture preferences that are not representative of the data. Again, the IV estimates do not change much from the main results.

Taken together, the results show that better-connected households contribute more to various public goods and also contribute overall. The positive effects are robust to several robustness checks.

1.5.4 Alternative specifications

This subsection considers alternative specifications to the main results.

Non-linear relationships

It is possible that there is a nonlinear relationship between contribution and network centrality. Bandiera and Rasul (2006) found an inverse-U relationship between the number of social links and the decision to adopt a new crop because of the learning effects and strategic delays. In the current context, the influence from each link may decline with additional links. It could also be that households with a very large network can assert influence on their links and do not contribute. I explore two different specifications that allow for non-linearity. Since there is only one IV for network centrality, the identification strategy cannot be applied to the non-linear specifications.[34] Investigating the possibilities for non-linear relationships using OLS estimates can still help interpret the main results.

I first formulate a nonlinear specification by adding a squared term of degree centrality to the OLS specification (in Panel B of Table 1.8). Table 1.11 shows that the coefficients for squared terms are negative, albeit small and statistically insignificant, for most public goods. The next specification (Table 1.12) allows for more flexibility by using a set of dummy variables for each level of degree centrality. There is some evidence suggesting nonlinear relationship. In particular, for water well contribution, having more than five links seems to have a stronger positive correlation. I test the non-linear models against the linear specifications using the likelihood ratio test. The results show that the two models are not statistically different. Overall, however, there is limited evidence to suggest non-linear relationships between centrality and public good contribution.

[34] One possibility is to use a squared term of household ethnic fraction as an additional instrument for a squared term of centrality. However, the first-stage F statistics do not pass the weak instrument test.

Peer effects framework

Within the existing literature, a number of studies have explored impacts of social networks under a peer effects framework which posits that an agent's behavior is influenced by peer behaviors. My paper focuses on the effect of network position (measured by centrality) rather than peer effects. For comparison purposes, I also report the results using a peer effects specification. I regress household public good contribution on the number of connected households that contribute, controlling for the total number of connections (degree centrality).

$$Contribute_i = \alpha_0 + \alpha_1 \sum_{j \neq i} g_{ij} Contribute_j + \alpha_2 \sum_{j \neq i} g_{ij} + \alpha_3 X_i + e_i$$

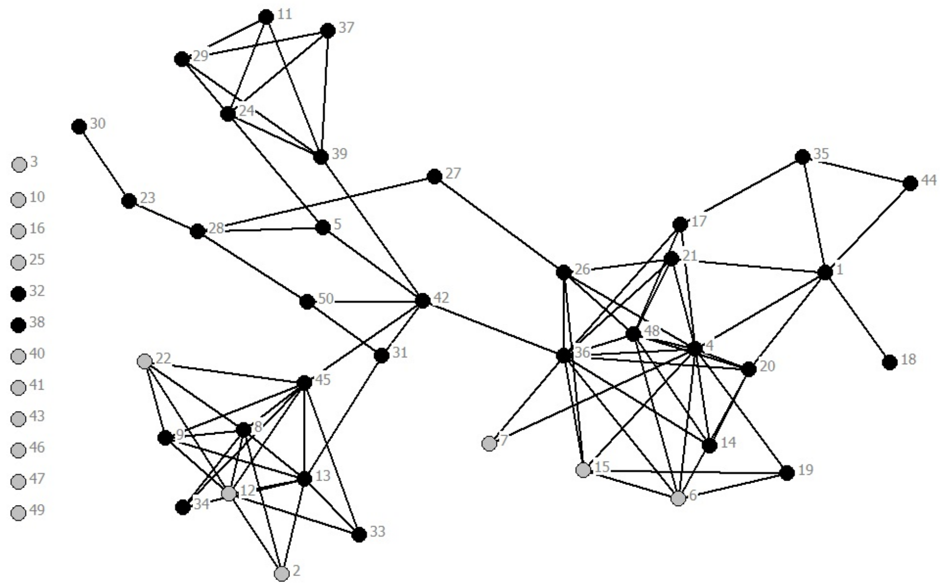
where $g_{ij} = 1$ if household i is connected to household j; $g_{ij} = 0$, otherwise. α_1 is the effect of having an additional link who contributes.

Table 1.13 shows that, in most cases, a household decision to contribute is positively correlated with its peers' behaviors, indicating possible strategic complementarities assumed in some of the models discussed in Section 2. The coefficients are significant only for roads and community buildings. Holding the total number of peers constant, an additional peer who contributes is associated with a 1% increase in likelihood to contribute to roads. The estimates however could be resulted from correlated unobserved factors such as preferences for public goods. In that case, the positive sign indicates that households with similar preferences are more likely to be connected. Identifying the peer effects is beyond the scope of this paper.

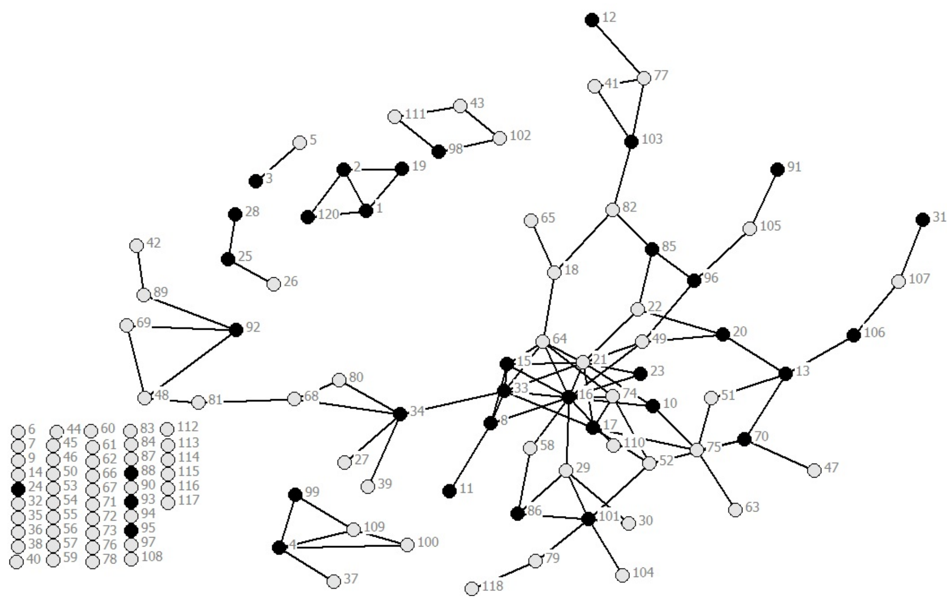
1.6 Conclusion

This paper employs social network analysis from the sociology literature to analyze economic decisions. I provide field-based evidence for the relationship between network position and household public good decisions, a relationship that has mostly been studied in the theory literature and the experimental setting. The results establish network centrality as a determinant increasing public good contribution. Given the limitations of the current data, I do not distinguish between the mechanisms driving this relationship. A natural next step would be to understand these distinctions better. Additionally, the descriptive results at the village level suggest a possibility that sparse networks may constitute a channel that causes the negative relationship between ethnic diversity and public goods. This observation can potentially explain some of the puzzling results in Esteban et al. (2012), in which countries such as China, Haiti, and undivided Korea experienced many cooperative problems despite low ethnic fragmentation. Future works can investigate these particular cases for evidence regarding their network structures. The findings also yield policy implications that exploiting information about existing social networks may be beneficial when implementing community-driven projects. For example, policy makers can target their effort to communities with more households with low centrality, because it is more difficult to sustain cooperation.

1.7 Figures



(a) Village ID: 4, contribution to water well construction



(b) Village ID: 11, contribution to public health

FIGURE 1.1: Examples of kin network graphs (black dot = contribute)

Notes: This figure presents network graphs of kin networks for two villages in the sample. Each node represents a household in the village and a line represents a kin relationship between the two connected households. The black color means the household contributes to the public good.

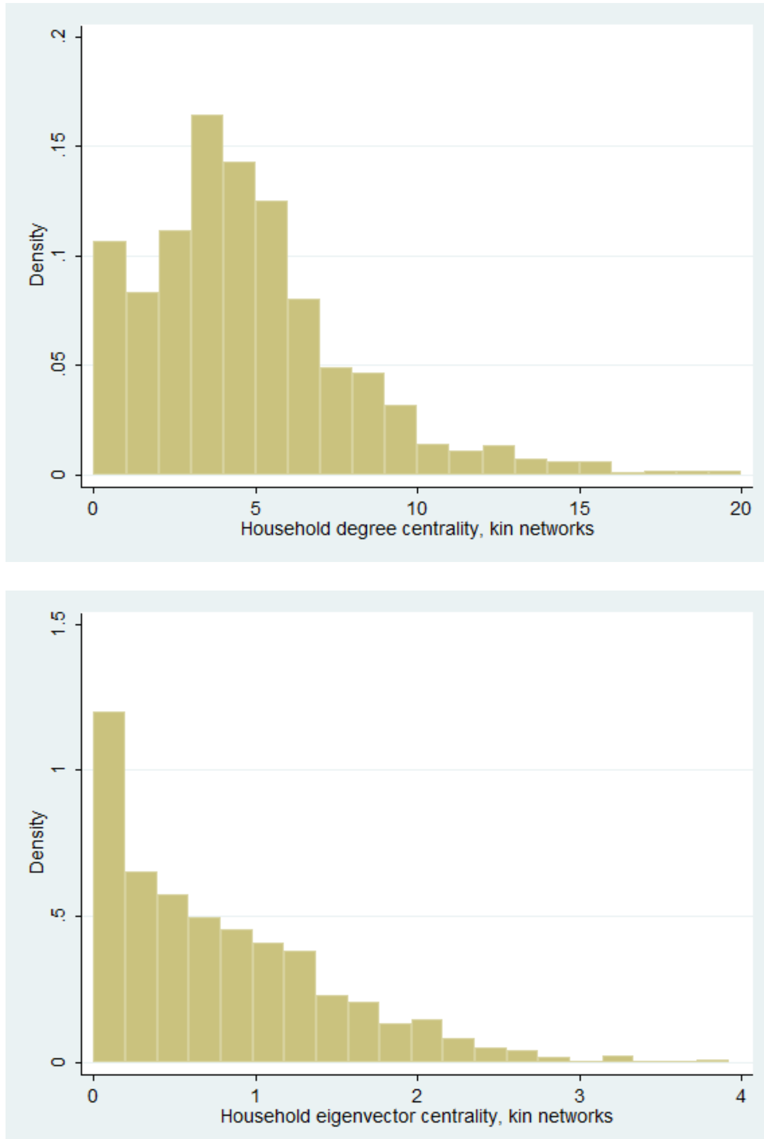


FIGURE 1.2: Distribution of centrality measures for kin networks

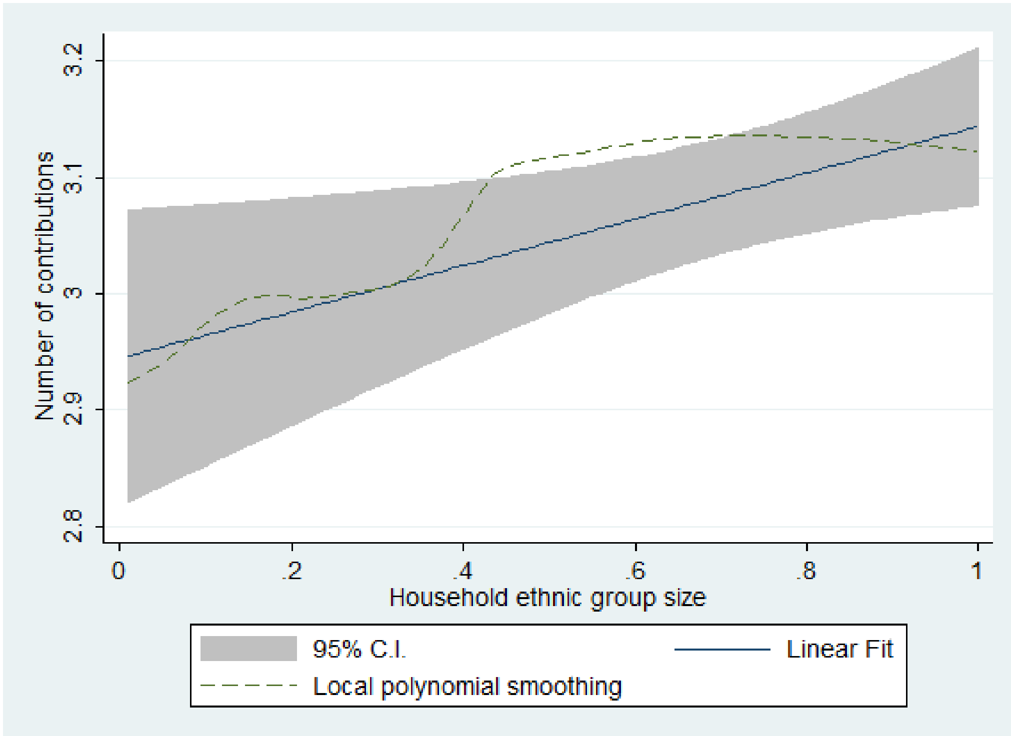


FIGURE 1.3: Reduced-form relationship between number of household contributions and household ethnic fraction

1.8 Tables

Table 1.1: Summary statistics at the village level

Variables	Mean	S.D.
A: Village characteristics		
# of households	54.31	28.10
Total population	599.08	210.95
Village area (km^2)	122651.40	127730.90
Log per capital income	14.74	0.58
Income Gini coefficient	0.30	0.10
Illiteracy rate	0.42	0.21
% Population without electricity	0.97	0.04
% Population without toilet	0.37	0.30
Ethnic diversity, Herfindahl index	0.31	0.21
Religion diversity, Herfindahl index	0.03	0.08
B: Outcome variables		
% contributing to road construction	0.56	0.29
% contributing to community building construction	0.60	0.33
% contributing to water well construction	0.45	0.27
% contributing to public health	0.58	0.31
% contributing to public education	0.36	0.30
Avarage total contribution	2.71	0.98
Number of observations	48	

Table 1.2: Summary statistics at the household level

Variable	Mean	S.D.
A: Household Characteristics		
Income (Gambian Dalasi, 28GD = 1USD in 2008)	29213.18	20506.27
Household size	12.80	8.79
Head age	51.97	16.15
Female head	0.06	0.23
Whether household head is illiterate	0.46	0.50
# Current sick members	0.24	0.49
# Household workers	4.55	3.72
Amount of lands owned (hectares)	8.29	20.50
Self-reported high quality land	0.05	0.21
Marital status1: unmarried	0.03	0.17
Marital status2: monogamous	0.46	0.50
Marital status3: polygamous	0.48	0.50
Farming	0.73	0.44
Muslim	0.99	0.09
Number of corrugated huts	1.31	1.37
Number of grass-roof huts	1.09	1.55
Self-report relative wealth 1 (first quantile)	0.45	0.50
Self-report relative wealth 2 (second quantile)	0.25	0.48
Self-report relative wealth 3 (third quantile)	0.16	0.36
Self-report relative wealth 4 (highest quantile)	0.04	0.20
Role: Alkalo (village chief)	0.02	0.14
Role: Village development committee	0.19	0.40
Role: Traditional healer	0.13	0.34
Role: Imam	0.02	0.13
Role: Marabout	0.02	0.14
Ethnicity 1: Mandinka	0.53	0.50
Ethnicity 2: Fula	0.22	0.41
Ethnicity 3: Wollof	0.10	0.30
Ethnicity 4: Jola	0.09	0.29
Ethnicity 5: Sarehuleh	0.01	0.08
Ethnicity 6: Sereer	0.04	0.20
Ethnicity 7: Manjago	0.01	0.09
Ethnicity 8: Others	0.01	0.08
B: Outcome Variables		
Contribute to public goods (Y/N)	0.97	0.16
Buliding road	0.62	0.49
Building community buildings	0.71	0.46
Public education	0.51	0.50
Public health	0.67	0.47
Public water source	0.44	0.50
Total contribution	3.10	1.34
Number of Observations:	2,112	

Table 1.3: Household network centrality and village network structure

Variables	Mean	S.D.
A: Village-level network characteristics		
Kin network, density	0.13	0.08
Number of Observations:	60	
B: Household-level network characteristics		
Kin network, degree centrality	4.16	3.03
Kin network, eigenvector centrality	0.79	0.71
C: External link (binary)		
External marriage link	0.85	0.36
External land-exchange link	0.16	0.37
External labor-exchange link	0.08	0.27
External input-exchange link	0.11	0.32
External credit-exchange link	0.15	0.36
Number of Observations:	2,112	

Notes: All networks are undirected and unweighted. Eigenvector centrality is normalized within a village. See Appendix A for formal definitions of network terminology.

Table 1.4: Comparing households of the same ethnic group with variation in ethnic fraction

	(1) Coeff.	(2) Std. Errors
Log income	-0.224	(0.925)
Household size	0.038	(0.086)
Head age	-0.042	(0.035)
Female head	-0.395	(1.990)
Illiterate head	1.484	(1.344)
# Sick members	-0.693	(0.798)
# household workers	0.066	(0.148)
Land owned (hectares)	0.021	(0.018)
Self-reported high-quality land	0.693	(1.417)
Role: Alkalo	1.762	(1.497)
Role: Village dev. committee	1.700	(1.180)
Role: Traditional healer	0.364	(1.204)
Role: Imam	4.060**	(1.630)
Role: Marabout	2.383	(1.630)
Farming	1.079	(1.195)
Muslim	1.917	(3.526)
Marital status1: unmarried	2.171	(3.431)
Marital status2: monogamous	0.318	(2.937)
Marital status3: polygamous	0.339	(2.789)
Relative wealth 2: second quantile	1.009	(1.024)
Relative wealth 3: third quantile	0.372	(1.174)
Relative wealth 4: highest quantile	-1.975	(1.812)
Agricultural income	-0.022	(0.101)
Number of corrugated huts	0.640	(0.419)
Number of grass-roof huts	-0.709	(0.484)
Observations	2,110	

Notes: This table shows that, within the same ethnicity, households from majority groups and minority groups are comparable in all but one characteristic. Robust standard errors clustered at the village level are reported in parentheses.
*** p<0.01, ** p<0.05, * p<0.1.

Table 1.5: Network structure and village-level contribution

VARIABLES	(1) Road construction	(2) Communal buildings	(3) Water well construction	(4) Public health	(5) Public education	(6) Total contribution
Panel A: OLS estimates without controlling for network density						
Ethnic diversity index	-0.032	-0.037	0.127	-0.165	-0.113	-0.549
	(0.169)	(0.181)	(0.151)	(0.167)	(0.189)	(0.529)
Panel B: OLS estimates controlling for network density						
Ethnic diversity index	0.012	-0.059	0.043	-0.064	-0.104	-0.313
	(0.167)	(0.200)	(0.152)	(0.153)	(0.193)	(0.494)
Network density	0.801*	-0.383	-1.496***	1.818***	0.163	4.253**
	(0.466)	(0.820)	(0.426)	(0.551)	(0.918)	(1.975)
Mean outcomes	0.53	0.61	0.45	0.58	0.39	2.69
Number of Obs	60 villages					

Notes: This table presents the OLS estimates from regressing aggregate public good contribution on village ethnic diversity (Herfindahl index) and network density. Each cell presents the coefficient of network density for each separate regression. All specifications include district fixed effects and the following control variables: number of households, total population, areas, log per capita income, income Gini coefficient, illiteracy rate, % of households with no electricity, religion diversity (Herfindahl index). The coefficients of village characteristics are not reported here, but are available upon request. Standard errors are in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.6: First stage regressions

	(1)	(2)
Endogenous Variable:	Degree cent.	Eigen. Cent.
IV:		
Household ethnic group size	3.955***	0.874***
	(0.506)	(0.090)
Constant	-1.260	-0.196
	(1.537)	(0.501)
First-stage F statistic	61.03	94.45
Observations	2,110	
R-squared	0.388	0.172
# villages	48	

Notes: This table shows the first stage regression using *household ethnic group size* to instrument for network centrality. The critical value for the F-statistics is 16.38 (Stock and Yogo (2005)). The coefficients of household characteristics are not reported here, but are available upon request. Robust standard errors clustered at the village level are reported in parentheses. The specification includes ethnic-group FE, village FE and the following control variables: log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, role as the village chief, role as a traditional healer, role as a village development committee, role as a imam, role as a marabout, religion, marital status, relative wealth, agricultural income, number of corrugated huts and number of grass-roof huts. *** p<0.01, ** p<0.05, * p<0.1.

Table 1.7: Correlates of centrality measures

	Degree cent.		Eigen. cent.	
	(1)	(2)	(3)	(4)
	Coeff.	Std. Errors	Coeff.	Std. Errors
Log income	-0.158	(0.135)	-0.038	(0.052)
Household size	0.024*	(0.013)	0.007**	(0.003)
Head age	0.017***	(0.004)	0.003***	(0.001)
Female head	-1.039***	(0.303)	-0.084	(0.099)
Illiterate head	-0.045	(0.148)	-0.006	(0.039)
# Sick members	-0.037	(0.142)	-0.067*	(0.034)
# household workers	0.032	(0.024)	0.004	(0.006)
Land owned (hectares)	0.012**	(0.005)	0.002***	(0.001)
Self-reported high-quality land	0.816**	(0.311)	0.089	(0.091)
Role: Alkalo	1.796***	(0.495)	0.389***	(0.107)
Role: Village dev. committee	0.735***	(0.181)	0.156***	(0.044)
Role: Traditional healer	-0.192	(0.191)	-0.089	(0.055)
Role: Imam	0.332	(0.415)	-0.046	(0.111)
Role: Marabout	0.581	(0.382)	0.091	(0.095)
Farming	0.266*	(0.153)	0.115**	(0.044)
Muslim	0.152	(0.689)	0.159	(0.141)
Marital status1: unmarried	0.324	(0.471)	0.101	(0.135)
Marital status2: monogamous	0.340	(0.400)	0.083	(0.105)
Marital status3: polygamous	0.620	(0.378)	0.154	(0.098)
Relative wealth 2: second quantile	-0.067	(0.149)	0.005	(0.046)
Relative wealth 3: third quantile	-0.206	(0.205)	-0.022	(0.063)
Relative wealth 4: highest quantile	0.227	(0.474)	0.273**	(0.128)
Agricultural income	0.012***	(0.002)	0.002***	(0.001)
Number of corrugated huts	0.489***	(0.081)	0.098***	(0.019)
Number of grass-roof huts	0.016	(0.056)	-0.009	(0.020)
Observations	2,110		2,110	

Notes: This table show correlations of the two network centrality measures with household characteristics. The high correlations with some variables indicate that network centrality is likely endogenously determined. Robust standard errors clustered at the village level are reported in parentheses.
** p<0.01, ** p<0.05, * p<0.1.

Table 1.8: OLS results

Outcome variable	(1) Road construc.	(2) Commu. buildings	(3) Water well construc.	(4) Public health	(5) Public educ.	(6) Total contrib.
Panel A: Correlations between network centrality and public good contribution						
Obs. = 2,205						
Degree	-0.000	0.006*	0.015***	-0.000	0.014***	0.020*
	(0.003)	(0.003)	(0.005)	(0.004)	(0.005)	(0.011)
Eigenvector Cent.	0.013	0.022*	0.060***	-0.014	0.073***	0.079*
	(0.016)	(0.013)	(0.016)	(0.017)	(0.017)	(0.044)
Household char.	no	no	no	no	no	no
Panel B: OLS controlling for household characteristics						
Obs. = 2,110						
Degree	-0.002	0.003	0.012**	-0.000	0.009*	0.010
	(0.004)	(0.004)	(0.005)	(0.004)	(0.005)	(0.011)
Eigenvector Cent.	0.011	0.009	0.053***	-0.017	0.062***	0.048
	(0.017)	(0.013)	(0.016)	(0.019)	(0.016)	(0.049)
Household char.	yes	yes	yes	yes	yes	yes
Mean of outcome	0.61	0.71	0.5	0.67	0.44	3.1

Notes: This table shows the OLS estimates from regressing public good contribution on household network centrality. Each cell presents the main coefficient of interest for each separate regression. All specifications include ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, role as the village chief, role as a traditional healer, role as a village development committee, role as a imam, role as a marabout, religion, marital status, relative wealth, agricultural income, number of corrugated huts and number of grass-roof huts. The coefficients of household characteristics are not reported here, but are available upon request. Robust standard errors clustered at the village level are reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.9: Main results: household ethnic fraction as an IV for kin network centrality

Outcome variable:	(1) Road construction	(2) Community buildings	(3) Water well construction	(4) Public health	(5) Public education	(6) Total contribution	First stage F-stat
Kin networks							
Degree cent.	0.006	0.016	0.028*	0.009	0.011	0.050**	61.03
	(0.012)	(0.013)	(0.014)	(0.009)	(0.011)	(0.023)	
Eigenvector cent.	0.028	0.071	0.127*	0.040	0.048	0.227**	94.45
	(0.056)	(0.056)	(0.069)	(0.040)	(0.050)	(0.097)	
Mean of outcome	0.61	0.71	0.5	0.67	0.44	3.1	
Number of observations			2,110				

Notes: This table shows the IV estimates using *household ethnic fraction* as an IV for network centrality. Each cell presents the main coefficient of interest for each separate regression. All specifications include ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, role as the village chief, role as a traditional healer, role as a village development committee, role as a imam, role as a marabout, religion, marital status, relative wealth, agricultural income, number of corrugated huts and number of grass-roof huts. The coefficients of household characteristics are not reported here, but are available upon request. Robust standard errors clustered at the village level are reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.10: Robustness checks

Panel A: Additional controls - external links and enumerator effects						
	(1)	(2)	(3)	(4)	(5)	(6)
Outcome variable:	Road construc.	Commu. buildings	Water well construc.	Public health	Public educ.	Total contrib.
Obs: 2,110						
Degree cent.	0.012	0.021	0.029*	0.010	0.012	0.064**
	(0.014)	(0.013)	(0.015)	(0.010)	(0.010)	(0.027)
Eigenvector cent.	0.054	0.096	0.133*	0.044	0.054	0.292**
	(0.061)	(0.057)	(0.074)	(0.044)	(0.045)	(0.110)
Mean of outcome	0.61	0.71	0.5	0.67	0.44	3.1
Panel B: Full sample						
Obs: 2,718						
Degree cent.	0.006	0.019*	0.031***	0.006	0.021**	0.058**
	(0.010)	(0.010)	(0.011)	(0.008)	(0.010)	(0.024)
Eigenvector cent.	0.025	0.083*	0.135***	0.026	0.092**	0.253**
	(0.044)	(0.044)	(0.049)	(0.037)	(0.042)	(0.100)
Mean of outcome	0.62	0.71	0.51	0.67	0.44	3.11
Panel C: Full sample excluding households with ethnic fraction >0.98 or <0.02						
Obs: 2,339						
Degree cent.	0.006	0.021**	0.031***	0.004	0.022**	0.056**
	(0.011)	(0.010)	(0.011)	(0.009)	(0.010)	(0.026)
Eigenvector cent.	0.026	0.091**	0.136***	0.017	0.094**	0.244**
	(0.048)	(0.044)	(0.049)	(0.041)	(0.042)	(0.107)
Mean of outcome	0.58	0.71	0.45	0.68	0.48	3.16

Notes: This table shows the 3 robustness checks as described in Section 5. Panel A reestimates the IV regressions controlling for an extensive set of external network variables and enumerator FE. Panel B expand the sample to the full sample as described in Section 4. Panel C restrict the sample to exclude households at the extreme tails of the ethnic fraction distribution. Each cell presents the main coefficient of interest for each separate regression. All specifications include ethnic-group FE, village FE, and the following characteristics: log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, role as the village chief, role as a traditional healer, role as a village development committee, role as a imam, role as a marabout, religion, marital status, relative wealth, agricultural income, number of corrugated huts and number of grass-roof huts. The coefficients of other variables are not reported here, but are available upon request. Robust standard errors clustered at the village level are reported in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.11: OLS regressions with squared terms

Outcome Variable	(1) Road construc.	(2) Community building	(3) Water well construc.	(4) Public health	(5) Public educ.	(6) Total contrib.
Panel A: Degree centrality						
Degree cent.	-0.002	0.012	0.016	0.004	0.005	0.024
	(0.010)	(0.010)	(0.010)	(0.008)	(0.011)	(0.024)
Degree cent. squared	0.000	-0.001	0.000	-0.001	0.000	-0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
Panel B: Eigenvector centrality						
Eigenvector cent.	-0.074	0.070	0.205	-0.054	0.198**	0.086
	(0.096)	(0.070)	(0.125)	(0.095)	(0.092)	(0.243)
Eigen. cent. squared	0.195	0.029	0.386	-0.047	-0.120	0.009
	(0.285)	(0.217)	(0.364)	(0.283)	(0.199)	(0.717)
Mean of outcome	0.58	0.70	0.48	0.66	0.48	3.11
Number of obs.			2,100			

Notes: This table shows the OLS estimates similar to those in Panel B of Table 1.8 but with a squared term of network centrality. This purpose of this estimation is to investigate the possibility of non-linear relationship between public good contribution and network centrality. Each cell presents the main coefficients of interest for each separate regression. All specifications include ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, role as the village chief, role as a traditional healer, role as a village development committee, role as a imam, role as a marabout, religion, marital status, relative wealth, agricultural income, number of corrugated huts and number of grass-roof huts. The coefficients of household characteristics are not reported here, but are available upon request. Robust standard errors clustered at the village level are reported in parentheses.
*** p<0.01, ** p<0.05, * p<0.1

Table 1.12: Nonlinear relationships between network degree centrality and public good contributions for combined networks

Outcome Variable	(1) Road construc.	(2) Community building	(3) Water well construc.	(4) Public health	(5) Public educ.	(6) Total contrib.
Degree: 1-2	-0.036	0.036	0.062	0.039	0.036	0.115
	(0.042)	(0.035)	(0.043)	(0.028)	(0.031)	(0.082)
Degree: 3-4	-0.022	0.051	0.060	0.037	-0.010	0.094
	(0.038)	(0.039)	(0.051)	(0.031)	(0.040)	(0.096)
Degree: 5-6	-0.029	0.055	0.093*	0.032	0.061	0.151
	(0.042)	(0.041)	(0.047)	(0.035)	(0.041)	(0.111)
Degree: 7+	-0.027	0.040	0.126**	0.023	0.074*	0.133
	(0.041)	(0.045)	(0.057)	(0.038)	(0.043)	(0.115)
R-squared	0.410	0.425	0.350	0.418	0.400	0.507
Observations			2,110			

Notes: This table shows the OLS estimates similar to those in Panel B of Table 1.8. Instead of degree centrality, I use categorical variables for different levels of degree centrality to investigate the possibility of non-linear relationship between public good contribution and network centrality. Each cell presents the main coefficients of interest for each separate regression. All specifications include ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, role as the village chief, role as a traditional healer, role as a village development committee, role as a imam, role as a marabout, religion, marital status, relative wealth, agricultural income, number of corrugated huts and number of grass-roof huts. The coefficients of household characteristics are not reported here, but are available upon request. Robust standard errors clustered at the village level are reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.13: Peer effects specifications

Outcome Variable	(1) Road construc.	(2) Commu. building	(3) Water well construc.	(4) Public health	(5) Public educ.	(6) Total contrib.
# Links who contribute	0.010*	0.005*	0.003	0.008	0.010	0.002
	(0.006)	(0.003)	(0.003)	(0.007)	(0.007)	(0.002)
# Links (Degree cent.)	-0.008	-0.000	-0.003	0.007	0.004	0.005
	(0.005)	(0.005)	(0.006)	(0.005)	(0.006)	(0.015)
Number of Obs.			2,100			

Notes: This table shows the OLS estimates from a peer-effects specification which regresses household contribution on its peers' average contribution. Each cell presents the main coefficient of interest for each separate regression. All specifications include ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, role as the village chief, role as a traditional healer, role as a village development committee, role as a imam, role as a marabout, religion, marital status, relative wealth, agricultural income, number of corrugated huts and number of grass-roof huts. The coefficients of household characteristics are not reported here, but are available upon request. Robust standard errors clustered at the village level are reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

1.9 Appendix

1.9.1 Network Terminology

This appendix describes formal definitions of network terminology. For definitions of other network terminology, see Jackson (2007).

A graph (N, G) consists of a set of nodes (in this setting, households), $N = \{1, ...n\}$, and a $n \times n$ binary matrix, G, where g_{ij} represents the relation between i and j. g_{ij} is equal to 1, if i and j are connected and is equal to 0 otherwise.

- **Degree centrality**

  Degree centrality refers to the number of nodes connected to a given node i in a graph.

  $$d_i(G) = \#\{j : g_{ij} = 1\} = \sum_j g_{ij}.$$

- **Bonacich centrality** (Bonacich (1987))

  Let $c_i(\alpha, \beta)$ denote the Bonacich centrality of a network, G. The β parameter reflects the degree to which a node's centrality score is a function of the scores of those to whom it is connected.

  $$c_i(\alpha, \beta) = \sum_j (\alpha + \beta c_j) g_{ij}$$

  In matrix notation,

  $$c(\alpha, \beta) = \alpha(I - \beta G)^{-1} G1$$

  where "1" is a column vector of ones and I is an identity matrix. For the matrix in the parentheses to be invertible, it must be that $\beta \in (-\frac{1}{\lambda_g}, \frac{1}{\lambda_g})$, where λ_g is the largest eigenvalue of G. The α parameter only affects the length of $c(\alpha, \beta)$. It is a normalizing factor generally selected so that the Euclidean norm of the vector equals the number of nodes in the network: $\sum_i c_i(\alpha, \beta)^2 = n$.

- **Eigenvector centrality** (Bonacich (1972))

  Let e_i denote the eigenvector centrality associated with a network, G. The centrality of a node is proportional to the sum of the centrality of its connected nodes:

  $$\lambda e_i = g_{i1} e_1 + g_{i2} e_2 + \ldots + g_{in} e_n = \sum_j g_{ij} e_j.$$

In matrix notion: $\lambda e = Ge$, where λ is a proportionality factor. e is an eigenvector of G and λ is its corresponding eigenvalue (the largest eigenvalue, λ_g, to ensure positive entries in e).

Another way to interpret the eigenvector centrality is to view it as the limiting case of the Bonacich centrality where the β-parameter approaches the inverse of the largest eigenvalue of G (see a discussion in Bonacich (2007)). The eigenvector centrality of a node is then a weighted sum of walks that that node has starting from it. A walk of length 1 is assigned a value $1/\lambda'_g$ where $\lambda'_g \to \lambda_g$. A walk of length 2 is $(1/\lambda'_g)^2$ and so on.

- **Network density**

  Network density measures how connected a network is relative to all possible links. n is the number of nodes in N.

$$\text{density} = \frac{\#\text{all links}}{\binom{n}{2}}.$$

- **Undirected and unweighted links**

  An undirected link refers to a link in which if i is linked to j, then j is linked to i. An unweighted link refers to a relationship that does not have any intensity level associated with. The value of links as represented in an adjacency matrix is either 0 or 1.

1.9.2 Economic-exchange networks

In addition to kin networks, the survey also contains information on economic interactions of households within the village. Two given households are considered to have an *economic-exchange link* if they had transacted on land, labor, credit, or production inputs such as tools and fertilizer within one year prior to the survey. I refer to a network of this relationship as an *economic-exchange network*.[35]

For economic-exchange networks, the IV, *household ethnic fraction*, is not strong enough to predict household centrality. Even though there is a positive relationship between having economic transactions and belonging to the same ethnic group, such correlation does not translate to enough predictive power, especially when conditioning on having a kin link (Arcand and Jaimovich (2012)). This weak correlation, on the other hand, also emphasizes the fact that ethnic identity is not very salient in the

[35] I combine all types of economic transactions, following Jaimovich (2011).

Gambia setting. Ethnicity does not influence with whom households choose to engage in economic activities. Nonetheless, the descriptive results can help understanding the bigger picture of how different types of network relationships influence public good decisions in comparisons to the main results. I use the same OLS specification analogous to that in Table 1.8:

$$Contribute_{ijk} = \gamma_0 + \gamma_1 Centrality_{ijk} + \gamma_2 X_{ijk} + ETH_j + \upsilon_k + \epsilon_{ijk} \quad (1.1)$$

Next, I combine both kin networks and economic-exchange networks. In this case, the same IV passes the weak instrument test. This stronger relationship is mainly driven by kin networks; therefore, the results must be interpreted as a supplement to the kin network results. Table 1.14 reports the economic-exchange network and combined network results. The signs of the centrality coefficients align wtih the results for kin networks.

However, only community building contribution has the strongest positive correlation with centrality. The IV results for combined networks show the same signs and significance patterns as the kin network results. This is expected mainly because kin networks comprise a larger portion of combined networks than economic exchange networks. To conclude, it seems that kin networks rather than economic exchange networks is the key factor influencing household decisions to contribute to public goods.

Table 1.14: Analysis of economic-exchange and combined network centrality

Outcome variable:	(1) Road construction	(2) Communal buildings	(3) Water wells construction	(4) Public health	(5) Public education	(6) Total contribution
Panel A: OLS estimates						
Econ networks						
Degree cent.	-0.002	0.006**	0.007**	0.001	0.010**	0.015*
	(0.003)	(0.002)	(0.004)	(0.003)	(0.004)	(0.009)
Eigenvector cent.	0.002	0.025**	0.026	0.005	0.052***	0.090**
	(0.013)	(0.010)	(0.019)	(0.011)	(0.015)	(0.035)
Combined networks						
Degree cent.	0.001	0.003*	0.008***	-0.003	0.005**	0.003
	(0.002)	(0.002)	(0.003)	(0.002)	(0.003)	(0.005)
Eigenvector cent.	0.026	0.018	0.079***	-0.012	0.092***	0.112*
	(0.022)	(0.016)	(0.026)	(0.023)	(0.020)	(0.061)
Panel B: IV estimates						
Combined networks						
Degree cent.	0.005	0.013	0.022*	0.007	0.008	0.040**
	(0.010)	(0.010)	(0.012)	(0.007)	(0.009)	(0.017)
Eigenvector cent.	0.039	0.099	0.175*	0.055	0.066	0.313**
	(0.077)	(0.081)	(0.095)	(0.055)	(0.064)	(0.124)
Mean of outcome	0.61	0.71	0.5	0.67	0.44	3.1

Notes: This table shows the OLS and 2SLS estimates similar to those in Panel B of Table 1.8 and Table 1.9. Instead of kin networks, I consider economic-exchange networks and combined networks here. The first stage regressions of the 2SLS results are reported in Table 1.6. Each cell presents the main coefficients of interest for each separate regression. All specifications include ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, role as the village chief, role as a traditional healer, role as a village development committee, role as a imam, role as a marabout, religion, marital status, relative wealth, agricultural income, number of corrugated huts and number of grass-roof huts. The coefficients of household characteristics are not reported here, but are available upon request. Robust standard errors clustered at the village level are reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

1.9.3 Additional Tables

Table 1.15: Reduced-form results

Outcome Vars	(1) Road construction	(2) Commu. buildings	(3) Water well construction	(4) Public health	(5) Public education	(6) Total contribution
Ethnic fraction	0.020	0.080	0.123*	0.026	0.034	0.187*
	(0.056)	(0.051)	(0.063)	(0.038)	(0.049)	(0.097)
Number of obs			2,083			

Notes: This table shows the reduced-form estimates from regressing public good contribution directly on the IV, household ethnic group size. Each column presents the coefficient of household ethnic fraction for each type of public good contribution. All specifications include ethnic-group FE and village FE. The set of household characteristics include log income, household size, age of household head, whether head is female, illiteracy, number of sick household members, number of workers, size of land owned, whether land is of high quality, role as the village chief, role as a traditional healer, role as a village development committee, role as a imam, role as a marabout, religion, marital status, relative wealth, agricultural income, number of corrugated huts and number of grass-roof huts. The coefficients of household characteristics are not reported here, but are available upon request. Robust standard errors clustered at the village level are reported in parentheses.
*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table 1.16: Ethnic composition at the national level across time

Year	Sources	Mandinka	Fula/Fulani	Wolof	Jola/Dyola	Sarehuleh	Serer	Manjago	Others
1983	The Gambia Resource page	36.6	17	13.2	9.4	7.4	2.3	1.6	
1992	The Gambia Resource page (1992 Priority survey)	33.4	17	12.8	n/a	n/a	n/a	n/a	n/a
1993	Alesina et al. 2003	34.1	16.2	12.6	9.2	7.7	n/a	n/a	n/a
2003	Gambia census 2003	36	22	14	11	8	3	2	4
2007	Current sample	48	21.8	9.6	7.5	2	4.8	0.9	5.4

Notes: The data in this table is based on the 1993 and 2003 censuses carried out by the Gambia Bureau of Statistics.

2

(Limited) Diffusion of Health-protecting Behaviors: Evidence from Non-beneficiaries of a Public Health Program in Orissa (India)

with Aprajit Mahajan and Alessandro Tarozzi

2.1 Introduction

Transmittable diseases such as malaria, yellow fever or intestinal worms, remain a heavy burden for public health in developing countries. In many cases, technological advances have created efficacious preventative measures. For instance, de-worming drugs are very effective at eliminating intestinal infections (Miguel and Kremer 2004) and insecticide treated nets can reduce considerably the burden of malaria (Lengeler 2004). The cost of such preventive technologies are very low for rich countries standards, but can be prohibitively expensive in low-income countries where neither individuals nor public health programs may have sufficient funding. A growing literature therefore studies the reasons of and possible solutions to the low uptake of health-protecting technologies in poor countries, see Holla and Kremer (2009) and

Dupas (2012) for recent reviews.

Given the low rates of adoption typically observed among the poor, the lack of experience with such potentially useful technology is often a factor that—together with budget constraints—further reduces demand. Several researchers have thus examined if social networks can facilitate the diffusion of health-protecting products. More generally, public health interventions that introduce such products on a large scale can generate important externalities through changes in disease environment, see Hawley et al. (2003) for the case of insecticide treated nets and malaria, or Miguel and Kremer (2004) for deworming drugs.

However, identifying network effects is hard. The main econometric problem lies in the endogeneity of social networks, as well-documented by a rich literature (see Manski 1993b, Brock and Durlauf 2001, Bramoullé et al. 2009 among others). To address this empirical problem, economists have used different strategies depending on the nature of the data. Some of the previous studies have used non-experimental data to tease out the network effects on the adoption of new agricultural technologies by making various identifying assumptions (Besley and Case 1994, Foster and Rosenzweig 1995, Munshi and Myaux 2006). Another key element in the estimation of network effects involves defining *which* social group constitutes the network. Some works use geographical or cultural proximity (Bertrand et al. 2000, Angelucci and Giorgi 2009, Dupas 2010). Others have argued that self-reported networks of friends and family are a better representation of social links (e.g. Bandiera and Rasul 2006, Conley and Udry 2010).

In this paper, we study the links between social networks and adoption of health-protecting technologies using data from a randomized controlled trial conducted in highly malarious areas of rural Orissa, India, between 2007 and 2009. The project was carried out in collaboration with the micro-lender Bharat Integrated Social Welfare Agency (BISWA) in 141 villages where BISWA lending activity was operational.

The main purpose of the field experiment was to evaluate the impact, relative to control conditions, on the take-up, usage and health effects of insecticide-treated bednets (ITNs) distributed either through micro-consumer loans or free of cost. Numerous studies have shown that high coverage and use rates of ITNs can significantly reduce malaria-related morbidity and mortality, see Lengeler (2004) for an extensive review of the evidence. Considerable evidence in particular shows that ITNs can be significantly more effective than bednets not treated with insecticide, because only the former nets can lead to externalities which benefit even individuals not sleeping under an ITN due to the reduction in mosquito density (Hawley et al. 2003, Killeen et al. 2007).

The results of the ITN distribution program on BISWA households' outcomes are described in detail in Tarozzi et al. (2011). The authors show that despite a substantial increase in ITN ownership and (self-reported) usage, especially in areas with free distribution, malaria indices did not improve during the study period. A key feature of the study was that the beneficiaries of the distribution programs were only 'BISWA households', that is, households where at least one individual was already affiliated to BISWA. On average, about 20% of the population in the 141 study villages had such affiliation, so that the majority of the local population was not directly affected by the program. Tarozzi et al. (2011) conjecture that the low coverage of the ITN distribution program was a leading cause for the lack of observed health benefits. If externalities that derive from high usage rates are key factors for the effectiveness of a health product, then public health programs that do not lead to such high coverage rates may lead to a waste of resources. Such concerns may be mitigated if within-community usage diffuses from beneficiaries to non-beneficiaries of distribution programs through network effects.

In this paper we analyze how the increased rates of ITN ownership and usage observed among BISWA households affected the same outcomes among non-beneficiaries, using data collected during the post-intervention survey, carried out in the winter of 2008-09. A key feature of our data is the availability of information on the number and type of social links between non-beneficiaries and a sample of BISWA households directly affected by the program. We then examine three specific questions. First, we estimate simple differences in outcomes between non-beneficiaries in control areas versus others residing in program areas where ITN ownership rates of BISWA households were exogenously increased by the program. Second, we examine if such differences where affected by the number and type of social ties between non-beneficiary and BISWA households. Although such ties are clearly endogenous, finding that the *interaction* between the ties and an (exogenous) program dummy matters for non-beneficiaries' outcomes would signal that spillover effects are present and likely mediated by the social links. Third, we estimate the effect of BISWA peers' behavior on the behavior of non-beneficiaries with instrumental variables, using program dummies as plausibly exogenous instruments for the (endogenous) peers' behavior.

Our paper contributes to a growing literature that uses experimental variation to estimate peer effects in the adoption of health-protecting technologies in developing countries. In a seminal paper, Kremer and Miguel (2007) showed that take-up of deworming drugs among schoolchildren in Kenya was *lower* among children with a larger fraction of peers exposed to a public health program of free treatment. They rationalize the result on the basis of the small private benefit and large positive externalities of the drug. In contrast, Dupas (2010) finds that experimental variation in the fraction of neighbors who received free or highly subsidized ITNs (a product with potentially high private returns) increases the likelihood of purchase. Kremer et al. (2011) find that random variation in the fraction of peers exposed to a point-of-use chlorine treatment for drinking water in Kenya had little impact on take-up. However, using an approach based on Graham (2008), they also find strong peer effects in the adoption of a point-of-collection water purification method based on a chlorine dispenser system. They explain the different results based on the public (point-of-collection) versus private (point-of-use) nature of the action required to adopt the two technologies.

The remainder of the paper proceeds as follows. Section 2.2 describes the experimental setup and descriptive statistics. We discuss the results on the spillover effects of the free ITN distribution on non-beneficiaries in Sections 2.3.1 and 2.3.2, while in Section 2.3.3 we estimate peer effects of ITN adoption and usage using instrumental variables. Section 2.4 concludes.

2.2 Data and Study Design

The results described in this paper are part of a broader evaluation of the cost effectiveness and health impacts of alternative mechanisms to deliver ITNs in poor areas of rural Orissa, India. Official figures pinpoint Orissa as the most highly malaria endemic state in the country (Kumar et al. 2007). The key element of the broader project was a large-scale cluster randomized controlled trial (RCT) designed to evaluate the uptake and impacts of insecticide-treated bednets (ITNs) through micro-consumer loans, as compared to free distribution and control conditions. The study was conducted in rural Orissa in 2007-09, in collaboration with Bharat Integrated Social Welfare Agency (BISWA), a micro-lender with an important presence in the study areas.

A baseline survey was completed in May-June 2007 with a sample of 1,844 households from 141 villages with BISWA presence. In all sampled households, at least one member was affiliated to BISWA, having joined a BISWA 'self-help group'. These are self-formed groups that can apply for micro-loans for which each member becomes jointly liable. We will refer to this sample of 1,844 households affiliated to BISWA as 'baseline' or 'BISWA' households.

After the baseline, villages were randomly assigned to one of three different experimental arms. In the fall of 2007, the project team carried out in all villages an information session about malaria and proper use of bed nets. In addition, in a first group of 47 villages ("Free" experimental arm), the team distributed free ITNs to all households with members affiliated to BISWA, along with the promise of two free insecticide retreatments at six month intervals. A second group of 47 villages ("MF") received offers to buy ITNs on credit, through micro-loans with a repayment term of one year at a 20% interest rate. The ITN offer price was not negligible, corresponding approximately to three to five times the local daily agricultural wage. Lastly, the control group received no other intervention beyond the information session.

About 96% of the households approached at baseline were then re-interviewed between December 2008 and May 2009, forming a panel of 1,768 households. In earlier work, Tarozzi et al. (2011) show that 52% of sample households purchased ITNs on credit in MF villages, although coverage in these locations remained significantly lower than what achieved with free distribution, where 96% of households received at least one ITN. Unexpectedly, neither micro-loans nor free distribution led to improvements in malaria and anemia prevalence.

A key element of the RCT was the focus on households with BISWA affiliation. Only these households were included in the delivery program, and all surveyed households were selected from lists of BISWA affiliates. In this paper, we focus instead on a supplementary sample added at the time of the post-intervention survey, in 2008-09. This additional sample was added for the purpose of studying impacts on non-beneficiaries.[1] The sampling frame was represented by publicly available census lists drafted as part of the Below the Poverty Line (BPL) census carried out in 2002

[1] Non-beneficiaries were only included in the post-intervention survey because the necessary funding was not available beforehand.

by the Ministry of Rural Development, Government of Orissa, with the purpose of identifying 'poor' households eligible to benefit from a number of welfare programs. Although our survey was carried out a few years after the census, preliminary observations in the field showed that the rosters remained overall reliable. In each study village, 10 new households were thus randomly chosen from the census lists, regardless of their BISWA membership. Interviews were then completed with a total of 1,425 new households, of which 1,153 were not affiliated with BISWA. In this paper, we analyze the behavior of these 1,153 non-BISWA households as a function of their indirect exposure to the ITN distribution program. In contrast, we do not use information about households with BISWA affiliation, because these would have likely been affected directly by the interventions.

In principle, the categorization of households as 'BISWA' or 'non-BISWA' could be problematic if affiliation with the micro-lender was endogenously affected by our ITN distribution programs. For instance, suppose that after the ITN distribution programs non-members with higher expected benefits from ITNs were more likely to join a BISWA self-help group, in the expectation of future ITN distribution programs. Then we would have likely observed a higher fraction of BISWA members among the new census-drawn households in Free or MF villages relative to control areas. However, the fraction of BISWA members in the population is similar in the three experimental arms (22.5, 20.5 and 20% in Control, Free and MF villages respectively) and the null of equality cannot be rejected at standard levels (p-value= .806). In the rest of the paper, we will thus assume that BISWA membership was exogenous.

2.2.1 Descriptive Statistics

The post-intervention survey was conducted using a uniform questionnaire for all interviewees, regardless of whether they were part of the panel or their BISWA affiliation. Enumerators recorded demographic, socio-economic, health and other indicators as well as detailed information on sleeping patterns, bednet ownership and use. In a key section of the questionnaire, we collected information about social ties between the respondent's household and each of the BISWA households interviewed in the same village before the intervention ('baseline households'). First, the respondent was asked if anyone in his/her household knew any member from each of the baseline households. Second, if the answer was yes, we recorded the frequency of the social contacts (daily, weekly, monthly, less than monthly). Third, we asked how often the social contacts involved conversations about health-related issues. Fourth, we asked "[w]hen you think about ways of protecting yourself and your family from malaria, do you take into consideration what persons in [panel household] do and what their opinions are?".

We use this information to construct three alternative measures of social links between non-BISWA households in the supplemental sample and baseline households from the same village. The first and most basic measure is the fraction of baseline households known to the respondent ('BISWA Network'). Next, we calculate the fraction of baseline households with whom the respondent's households interacted at least once a week ('Close BISWA Network') and finally the fraction whose opinions about ways of protecting oneself from malaria were taken into account ('Influential BISWA Network'). Previous empirical works on social networks have used a similar concepts, by directly asking about common conversation topics to identify sources of information (Kremer et al. 2011, Conley and Udry 2010).

We chose to use fractions instead of numbers because the number of baseline households are not always the same across villages. Suppose that a respondent reports links to, say, 10 baseline households. In a village where 15 baseline household were interviewed this would indicate links with an estimated 2/3 of BISWA households, while in a village where only 10 were interviewed we would estimate that a link exists with all BISWA household in the village, thereby indicating a stronger potential indirect exposure to the program. Recall that all households in the baseline survey were randomly chosen from a list of BISWA members within the village, whereas the new households were randomized from census list of that village. Therefore, the network measures as described above are unbiased estimators of the true fractions we would have observed if we had information about links to *all* BISWA households in the village.[2]

In Figure 2.1 we show the histogram of BISWA Network. The average network size of the entire sample was 0.64 so that, on average, non-BISWA households knew about two-thirds of BISWA households from the same village. The majority of households (68%) knew more than half of the BISWA links within their villages, with 260 households (23%) knowing everyone (BISWA Network = 1). Very few households had no or very few links with the baseline sample. In Figure 2.2 we show the histogram of Close BISWA Network. A comparison with Figure 2.1 shows that frequent interactions were not the rule, and on average close links only existed with less than half (44%) of baseline households. In total, 691 respondents (60%) reported frequent interactions with only half or less of the baseline BISWA households. Finally, Figure 2.3 shows the histogram of Influential BISWA Network. This looks overall similar to the histogram from close links, although there is more mass both on the bottom and the top of the distribution.

To further investigate the pattern of social networks in the sample, we explore the association between the network measures and a number of household characteristics. We estimate simple OLS regressions using the measures of social ties with BISWA households as dependent variable. All estimates also include village fixed effects,

[2] Note, however, that a corollary of this way of estimating links to BISWA households is that each BISWA Network variable is measured with error, and the error will be correlated with the fraction of households affiliated to BISWA in each village. For instance, in a village with only 10 BISWA households, all of them would have been interviewed, and BISWA Network would be estimated with no error. But in a village with 50 BISWA households, only 15 of them would have been included in the baseline sample, thereby increasing the measurement error of the BISWA Network variable. We have ignored these considerations so far, although they likely deserve further scrutiny.

because the network variable may partly reflect the fact that some villages are smaller in size and therefore have a closer-knit community. The results, in Table 2.1, show that the network measures are overall very weakly correlated with almost all of the socio-economic and demographic indicators included as regressors. The clear exception is the scheduled caste/scheduled tribe (SCST) dummy, whose coefficient is systematically significant at the 5% or lower level, and relatively large in magnitude. Households that belong to SCST are on average linked to 5-6 percentage points more BISWA households than non-SCST ones. This is perhaps not surprising, given that a large fraction of BISWA affiliates belong to SCST social groups.

Tarozzi et al. (2011) showed that the characteristics of BISWA households were overall balanced across arms. Here we look at cross-arm balance in selected summary statistics for the supplementary sample of non-BISWA households see. Because we do not have baseline data for these households, all statistics are derived from the post-intervention survey. We focus on household characteristics unlikely to have changed with the interventions. The results in Table 2.2 show that sample households are on average large and poor. Only 40% of sample households have access to electricity, and the average monthly expenditure per head is 692 Rupees, about 50 USD using the most recent parity purchasing power exchange rate (World Bank (2008)). When we carry out tests of equality of means across experimental arms, the null is never rejected at standard levels. This is also true for each of the three Network variables, although links to BISWA households appear to be slightly more infrequent in Free and above all MF communities. In panel B, we also look at village-level characteristics, using data from the 2001 Census of India. Villages in Free groups are larger that in the other two arms in terms of area and populations, but the differences are not statistically significant.

As documented in Tarozzi et al. (2011), bednets were already available in the study areas prior to the intervention, but treatment with insecticide was rare. About two-thirds of baseline BISWA households owned at least one bed net, but less than 10% owned at least one treated net. Almost all of these bed nets were purchased from the market, at a median price of 70 Rupees, about 1.5 times the typical daily wage for agricultural labor in the area. Free distribution of bednets from Government or NGO-driven public health program was very rare, outside of our intervention. This was consistent with our sampling frame which, to avoid contamination, excluded areas where such distribution programs had been or were expected to be conducted in the foreseeable future. Tarozzi et al. (2011) show that no such contamination appeared to have taken place during the duration of the evaluation.

2.2.2 Outcomes

Throughout the paper we focus on different indicators of bednet adoption, using data on ownership, purchases and usage. We look separately at ITNs and 'any net', where the latter includes all bednets regardless of treatment status. Information on bednets purchases and ownership was collected as follows. First, the respondent was asked to list all 'sleeping spaces' (indoors or outdoors) used by members during the previous night. We then recorded who slept in each space, and whether the space was protected by a bednet. If the answer to the latter question was yes, we recorded when the bednet was acquired, from which source and at which price, and whether and when the bednet had been treated with insecticide. We count a bednet as an ITN if it had been treated with insecticide up to six months before the interview. Standard bednets need to be periodically re-treated with chemicals in order to retain their insecticidal power. The frequency of the re-treatment depends on the type and concentration of the chemical used, but we choose six month because such was the appropriate time interval given the specifics of the insecticide used with the project nets (see Tarozzi et al. 2011 for details).[3]

[3] Unlike the standard bednets used in our project, Long-lasting insecticidal nets (LLIN) do not require period re-treatment with chemical. Such nets are becoming more common in public health

The focus on last-night usage reduces the possibility of recall error, although it is admittedly a noisy measure of regular usage.[4] Next, we asked if the household owned any other bednet besides those used the previous night, and if so we collected the same information detailed above about each additional bednet.[5] We categorize a bednet (or ITN) as 'recently acquired' (from any source) if it had been with the household for less than 18 months. Such nets were thus likely acquired after our interventions.

2.3 Results

We organize the results in three parts. First, we examine simple differences in mean outcomes between experimental arms. Second, we estimate the association between outcomes and the alternative measures of links to BISWA households across arms, indicating the likely presence of spillovers from BISWA to non-beneficiary households. Finally, we re-visit the standard problem of peer effects using instrumental variable estimation, making use of the exogenous variation in peers' outcomes that derive from the experimental setting.

2.3.1 Differences in Means

We first look at the simple cross-arm differences in outcomes. Recall that we are examining the behavior of non-BISWA households, none of whom was targeted by

programs, and their use is recommended by the World Health Organization. However, they remain rarely available in local markets in developing countries.

[4] Information on bednet usage was also independently collected in the household roster. For each member, we recorded whether the member had slept under a bednet the night before the interview, and whether the net had been treated in the previous six months. Data on usage from the two alternative sources are very highly correlated and so the results are substantively the same using either source. The similarity of the results is also reassuring because it reduces the likely extent of reporting error.

[5] Survey enumerators also asked permission to see the bednets. This allowed them to verify the presence of the bednets, their state of maintenance, and whether the net was from our distribution program (BISWA nets were clearly labeled and easily identified).

our distribution programs. Recall also that observed characteristics unlikely to be influenced by the interventions appeared to be balanced across the different arms, and that we maintain the assumption that BISWA affiliation was not affected by our intervention. Under these conditions, in the absence of any kind of spillovers to non-beneficiaries we should observe similar bednet ownership and usage rates in Free and MF villages relative to Control areas. On the other hand, there are at least four channels through which the interventions could have impacted bednet ownership and usage among non-beneficiary households. First, non-BISWA households may have been exposed directly to the short information session on malaria and bednet that took place in the fall of 2007. Second, behavior may have been affected later on through imitation or learning, mechanisms that we will explore in more details in Section 2.3.2. Third, some ITNs may have been transferred from BISWA households to non-beneficiaries as gifts or through sales. Fourth, the frequency of bednet usage may have affected by community-wide changes in the local mosquito population and malaria prevalence caused by our programs of ITN distribution, especially in areas where a large number of ITNs were delivered for free to all BISWA affiliates.

The results, in Table 2.3, show that for all but two outcomes the null of equal means cannot be rejected at standard levels. About one every three households acquired at least one bednet during the previous 18 months, while only one in ten acquired nets that had been recently treated with insecticide. The fraction of individuals who slept under a bednets or an ITN was slightly higher in treatment versus control areas, but the null of equality is never rejected at standard levels. In particular, less than 10% of individuals slept under a treated net (3.9% in control areas, 4.3% in Free and 4.7% in MF villages). The null of equality is rejected at the 10% (but not at the 5%) level only for the outcomes that measure recently acquired ITNs. While 9% of non-BISWA households acquired any ITNs in the last year and a half in control areas and 7% did in Free areas, the fraction was much higher at 14.2%

in MF villages.[6] Transfers of BISWA nets from baseline households to non-BISWA households via reselling or donations are unlikely to be the cause of such differences, because most of the recently acquired ITNs were reported to be purchased from the market. In fact, only seven households owned BISWA nets.

In the last two rows of Table 2.3 we also examine difference in malaria prevalence and hemoglobin levels (another key health indicator often associated to malaria). Both indices were measured through blood tests conducted with rapid diagnostic tests that delivered results within minutes, directly in the field (see Tarozzi et al. 2011 for details). All individuals of age below 10 or between 18 and 40 were targeted for testing.[7] Overall, 2,345 individuals of age from 876 households were tested for malaria, while hemoglobin was measured for 2,362 individuals from 881 households. The results show that both health indicators were very similar across treatment groups, and the null of equality is not rejected at standard levels. Malaria prevalence was very high, with about 20% of individuals testing positive. Hemoglobin level was 11.5 grams per deciliter (g/dl) of blood on average. This is low, given that 11 g/dl is sometimes taken as a threshold below which individuals are consider anemic (Thomas et al. 2006).

The remarkable similarity of malaria indices across arms is a strong indicator that the local epidemiological environment was not affected by the interventions. This is consistent with earlier studies that suggest that community protective effects of ITNs only emerge when 60% or more of sleeping spaces are covered. Tarozzi et al. (2011) show that such high coverage rates were never achieved by our interventions, which only targeted BISWA households. Indeed, the authors suggest that this may have

[6] The difference remains large and significant even after dropping two MF villages where BISWA households in the sample purchased an unusually large number of ITNs for resale.

[7] Thirty-two percent of individuals were not tested, with the proportion seven percentage points higher in Free communities, and two percentage points higher in MF villages. The joint null of equal testing rate is not rejected at standard levels (p-value= 0.1113), although the null of equality between Free and Control areas is rejected (p-value= 0.042).

been a key factor in explaining the lack of impacts of the ITN distribution programs on health indices observed even among the beneficiary households.

2.3.2 Heterogeneous Impacts as a Function of Links to Beneficiaries

We have shown that mean outcomes in Control villages appear to be very close to those observed in Free villages. In this section, we explore whether such aggregate results actually mask the existence of spillovers for non-beneficiary households with tighter links to BISWA members. Among BISWA members, Tarozzi et al. (2011) show that increases in ITN ownership and usage in Free and MF versus Control communities were substantively and statistically significant. At the time of the post-intervention survey (when data on the supplemental sample were collected as well), BISWA households owned on average 1.9 bednets in Control areas, 2.5 in MF villages and 3.4 with Free distribution. The random assignment of villages into experimental arms generated thus exogenous variation among non-BISWA households in their exposure to information about bed nets and insecticide treatment, through their links with BISWA households.

For a given outcome Y_{vi} for household i from village v, we estimate models such as the following:

$$Y_{vi} = \alpha \text{Network}_{vi} + \beta \text{Network}_{vi} \times \text{Free}_v + \gamma' X_{vi} + F_v + \epsilon_{vi}, \tag{2.1}$$

where Network_{vi} is one of the measures of links to a 'BISWA Network' described in Section 2.2.2, $Free_v$ denotes the treatment status, F_v is a village fixed effect, and X_i is a vector of household characteristics unlikely to have changed as a consequence of the intervention. Such control variables include household size, number of children under 5 years old, fraction of males in the household, number of household members who completed some schooling, number of rooms in the dwelling, access to electricity, and whether the household belongs to scheduled tribe/caste. We estimate all models using Ordinary Least Squares.[8] We calculate standard errors allowing for intra-village correlation of residuals.

Note that in estimating model (2.3.2) we only include observations from Control and Free villages. In the latter communities, virtually all sample BISWA households received ITNs during our intervention. The number of nets received was a function of the demographic composition of the households (with a ceiling of four ITNs per household) and was therefore exogenously determined by our field team. In contrast, demand for ITNs in MF villages was endogenously determined by BISWA members. In these latter communities, non-BISWA households with stronger links to program beneficiaries were then not necessarily 'exposed' to more nets, and this would complicate the interpretation of the results.

In the absence of any spillover to non-beneficiary households, we would expect $\hat{\beta}$ to be close to zero. This coefficient can be interpreted as the differential impact of the free ITN distribution on outcomes of non-BISWA households, as a function of their links to the beneficiaries of the program. A key limitation of this approach is that social links with BISWA households are clearly not exogenous. Individuals with stronger or more numerous ties are likely to be systematically different from others with smaller networks. Indeed, in Table 2.1 we have shown that some household characteristics (in particular SCST status) predict the size of the 'BISWA Network'. In addition, being part of a large network may be correlated with unobserved characteristics such as sociability and open-mindedness which, in turn, are likely to influence the propensity to adopt a new technology. In principle, for instance, large estimates of $\hat{\beta}$ could be due not to peer effects, but to more connected households having been more likely to be present during the bednet treatment with insecticide that often took place in public areas. In such case, any impact on bednet usage or treatment

[8] The linear probability models yield similar results compared to probit and logit models. These are available upon request.

rates would have been due to direct exposure to the implementation of the program rather than to any social dynamic.

While keeping these caveats in mind, we first estimate a specification where social links are measured by 'BISWA Network', defined as the fraction of the baseline BISWA households known to the non-beneficiary household i (see Section 2.2.1 for details). We look at four different outcomes: a dummy for households that acquired at least one ITN from any source in the 18 months before the interview; a similar dummy defined for nets regardless of treatment status; the fraction of household members that slept under an ITN the night before the interview, and the fraction who slept protected by a bednet regardless of treatment. The results are in Table 2.4. All estimates of α, which measures the association between the network variable and the outcome in control areas, are small and not significant at standard levels.

The estimate of β for recently acquired ITNs is close to zero and not significant, while the estimate for all acquired bednets is relative large (0.183) but again not significant at standard levels. In contrast, usage of both ITNs and any bednets appears to be more common for households with more social ties with BISWA affiliates. The estimates implies that, in non-beneficiary households in areas with Free distribution of ITNs, knowing all BISWA members in the baseline sample as opposed to none increases the fraction of individuals having slept under ITNs by 7 percentage points, and the fraction who slept protected by a net regardless of treatment by 16 percentage points. Both coefficients are significant, although only at the 10% level. These estimates are substantively large, given that the overall fraction of individuals who used bednets (treated or not) was well below 10% (see table 2.3).

Next, we explore whether the type of the social interactions plays a role in how networks affected bednet-related behavior. Because closer peers may be more influential in household decisions, the impacts of the intervention channeled through close peers may be stronger. We do not find much support for this hypothesis. When we use the two alternative measures ‘Close BISWA Network’ and ‘Influential BISWA Network’ described in Section 2.2.1, the results show weak evidence of network effects, see Table 2.5. The only outcome where β is significant (and only at the 10% level) is the dummy for recently acquired bednets. A household acquainted with all baseline BISWA households, and who is influenced by their viewpoints on malaria matters, has a 21 percentage point higher probability of having recently acquired at least one net relative to another household with no such social ties. More generally, the estimated impacts on net acquisition and usage are small and not significant for ITNs, while they are substantively large (but not significant, with the exception indicated above) for all nets regardless of treatment status.

In sum, we find only weak evidence supporting the view that the large increase in ITN ownership and usage observed among BISWA households who received nets free of cost was transmitted to non-beneficiaries through social ties. One simple explanation for this finding is that, given the absence of health benefits even among beneficiaries, and even in the presence of strong social ties between the two groups of households, non-beneficiaries simply did not have sufficient incentives to adopt a technology that, after all, was not protecting health effectively among their peers.

2.3.3 Instrumental Variable Estimation of Peer Effects

The previous section described models that aimed at identifying the spillover effects of the free ITN distribution on the behavior of non-beneficiaries, spillovers that may have taken place through social networks. An alternative identification strategy also allows to identify directly the impact of BISWA social contacts’ behavior on the choices of non-beneficiaries. Formally, we are interested in estimating the slope β_1 in the following simple model:

$$Y_{vi} = \beta_0 + \beta_1 X_{g,vi} + \epsilon_{vi}, \tag{2.2}$$

where, like before, Y_{vi} is a given outcome for household i in village v, and $X_{g,vi}$ is a measure of malaria-related behavior among i's social links, where social links are defined in one of the alternative ways described in Section 2.2.1. Given the likely endogenous sorting of individuals into social groups, OLS estimates would likely lead to positive and significant estimates for β_1, but such estimates would not be consistent for the true causal impact of $X_{g,vi}$ on the behavior of non-beneficiaries. However, the exogenous variation in the behavior of social links due to the randomized interventions provides a useful framework to estimate peer effects. Specifically, we can use the randomly assignment treatment status as an instrument for the endogenous behavior of peers.

We then use data from all three experimental arms and use the two treatment dummies MF_v and $Free_v$ as instruments for the endogenous variable $X_{g,vi}$. The two dummies are defined as $MF_v = 1$ for individuals in villages where ITNs were offered for sale on credit, and $Free_v = 1$ in villages where ITNs were delivered free of cost. Given that $X_{g,vi}$ will be an index of bednet ownership or usage, and given that such outcomes were significantly affected by the interventions, the instruments should be very strongly correlated with the endogenous variable. In contrast, instrument exogeneity is more demanding, because it requires that the only link between i's behavior (measured by Y_{vi}) and the treatment dummies passes through the behavior of the social links. A first reason why such assumption could fail is if the large increase in the fraction of village population protected by ITNs led to a reduction in the village-wide malaria prevalence. However, the results in Table 2.3 show barely any difference in malaria indices by experimental arms. The assumption could also fail if peers effects also work (as they are likely to) through indirect links. In other words, $X_{g,vi}$ measures only behavior among BISWA households included in the village-specific sample at baseline, but non-beneficiaries may have also been affected by the behavior of BISWA households not included in the sample, or by the choices made by others with social ties to BISWA households.[9] Although these caveats must be kept in mind, the availability of two instruments at least allows us to test their exogeneity by carrying out standard tests of overidentification.

The estimates are shown in Table 2.6. As expected, all OLS estimates are positive, although they are only large and significant when $X_{g,vi}$ is the average number of all bednets per person owned by the BISWA social links. However, when we estimate equation (2.3.3) using two-stage least squares $\hat{\beta}_1$ is always become smaller (in some cases even changing sign) and it is never significant at standard levels. As expected, there is no weak instrument problem, and the first stage F test is larger than 30 in both models. The overidentification tests also provide overall support to the hypothesis of instrument exogeneity, although the null is rejected at the 10% level when the outcome is a dummy for recently acquired ITNs. Overall, the results suggest that peer effects were not important elements for the diffusion of bednet usage in the sample.

2.4 Discussion and Conclusions

In this paper we have described evidence of some limited diffusion of bednet acquisition and usage from beneficiaries of an ITN distribution program in rural Orissa, India, to households that did not receive bednets during the intervention. Identification of such network effects hinged on the change in ITN adoption among the beneficiaries of a program of bednet distribution that was carried out in a randomly selected subset of the 141 study villages.

On the one hand, we have shown that, on average, there were very limited spillovers. On the other hand, we find that bednet usage was substantively and

[9] Network effects due to social ties with beneficiaries from other villages are instead very unlikely, because the study villages were spread over five different districts, so that most villages were geographically far apart.

significantly associated with some (but not all) measures of social links between non-beneficiaries and beneficiaries. This provides some evidence of network effects in the adoption of a health product that has potentially high protective power against malaria risk. However, given the endogeneity of social links, we cannot exclude that such associations were at least in part due to indirect effects of the programs mediated by channels different from, but correlated with, the number of social links between beneficiaries and non-beneficiaries.

In interpreting the results, it is useful to recall that the study area was very broad, covering 141 villages in five different districts. At the same time, the ITN distribution program was only conducted in communities where the micro-lender BISWA was operating at the time of the baseline survey. The external validity of our results should therefore be evaluated with caution. With this caveat, our results should be a useful contribution to a growing literature that evaluates the diffusion of health-protecting products through social networks in developing countries. Gauging the extent of such diffusion is particularly important in settings where public health programs only cover a fraction of the population at risk, and when coverage rates are a key element for the effectiveness of the program. This can be crucial, given the important role of externalities in fighting several transmittable diseases such as malaria or other insect-borne diseases, or intestinal worms.

In our context, the limited diffusion in the adoption of ITNs in a highly malarious area of rural Orissa may have been due to the overall absence of health benefits among the primary beneficiaries of the ITN distribution program. This may have been due to the low fraction of beneficiaries in the population, coupled with perhaps irregular usage of ITNs. These factors may have limited the effectiveness of ITNs, which has been otherwise convincingly documented in controlled conditions in the field. In contrast, a different RCT carried out in Kenya, Dupas (2010) found that the demand for bednets increased when a randomly determined higher fraction of social links had adopted the product. Although the different study area likely justifies results different from ours, Dupas (2010) describes how a large fraction of bednet users perceived a reduction in malaria risk, as well as little discomfort in using the nets supplied through the project. The absence of clear health benefits in our context is a likely key reason for the limited diffusion of ITN usage in our study area. If so, free or heavily subsidized ITNs extended to the whole population may have been the only way to reduce the malaria burden through this potentially important health product.

2.5 Figures

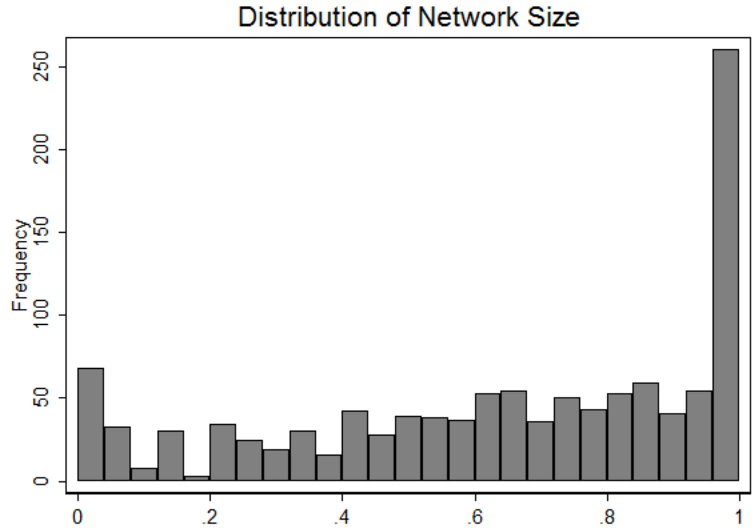


FIGURE 2.1: Distribution of BISWA Network

Data from winter 2008-09 from a supplemental sample of 1,153 households not affiliated to BISWA. BISWA Network is calculated as the fraction of baseline households from the same village known to the respondent's household. The overall sample mean is 0.64.

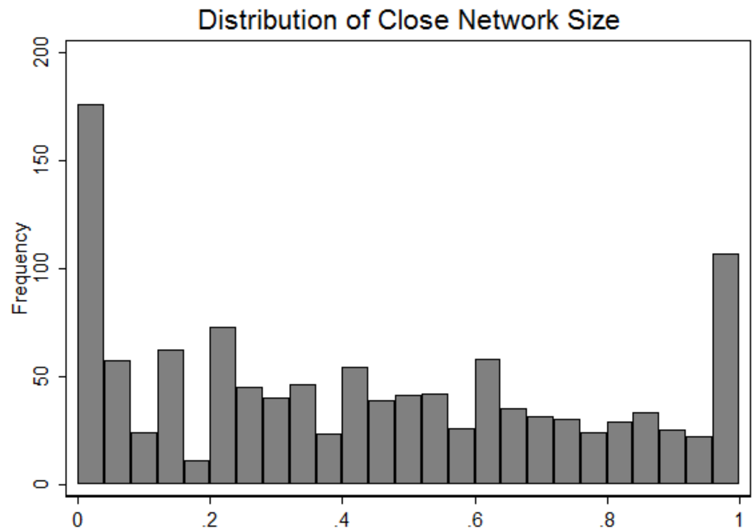


Figure 2.2: Distribution of Close BISWA Network

Data from winter 2008-09 from a supplemental sample of 1,153 households not affiliated to BISWA. Close BISWA Network is defined as the fraction of baseline households from the same village with whom the respondent's households interacts at least once a week. The overall sample mean is 0.44.

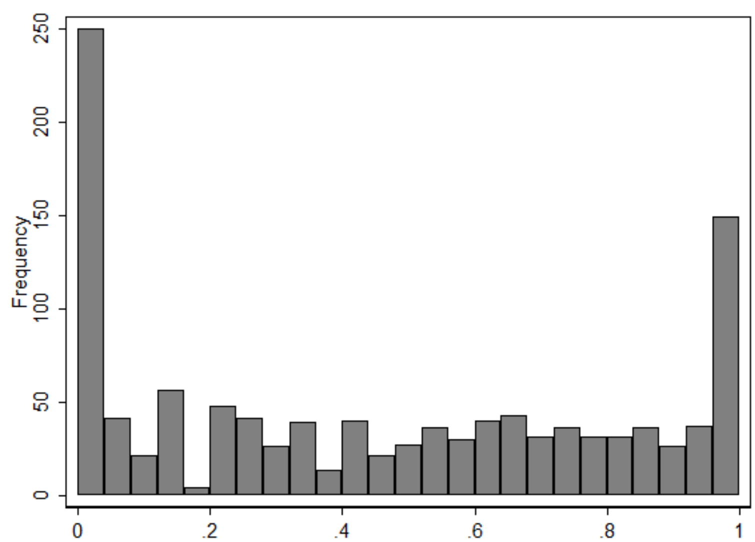


Figure 2.3: Distribution of Influential BISWA Network

Data from winter 2008-09 from a supplemental sample of 1,153 households not affiliated to BISWA. Influential BISWA Network is defined as the fraction of baseline households from the same village whose opinions about ways of protecting oneself from malaria are taken into account by the respondent's household. The overall sample mean is 0.46.

2.6 Tables

Table 2.1: Determinants of BISWA Network Size

	(1) BISWA Network	(2) Close BISWA Network	(3) Influential BISWA Network
Log per capita expenditure	0.005	0.013	0.015
	(0.015)	(0.017)	(0.015)
Household size	0.003	-0.005	0.016*
	(0.007)	(0.007)	(0.009)
Fraction of male	-0.031	-0.035	0.077
	(0.047)	(0.058)	(0.052)
#Children under 5	-0.022	0.006	-0.025
	(0.013)	(0.014)	(0.016)
#Members completed some schooling	0.007	0.010	-0.006
	(0.008)	(0.008)	(0.008)
#Rooms in the dwelling	0.010**	0.001	0.004
	(0.004)	(0.006)	(0.006)
Access to electricity	0.016	0.004	0.022
	(0.020)	(0.023)	(0.023)
Scheduled tribe/caste	0.048**	0.060***	0.055**
	(0.021)	(0.021)	(0.027)
Constant	0.520***	0.303***	0.207**
	(0.095)	(0.110)	(0.103)
Village Fixed Effects	Yes	Yes	Yes
Observations	1,150	1,150	1,150
R-squared	0.020	0.005	0.014

Data from winter 2008-09 from a supplemental sample of 1,153 households non affiliated to BISWA. Robust standard errors clustered at village level are reported in parentheses. BISWA Network is calculated as the fraction of baseline households from the same village known to the respondent's households. Close BISWA Network is defined as the fraction of baseline households with whom the respondent's households interacts at least once a week. Influential BISWA Network is defined as the fraction of baseline households from the same village whose opinions about ways of protecting oneself from malaria are taken into account by the respondent's household.
*** p< 0.01, ** $p < 0.05$, * $p < 0.1$.

Table 2.2: Descriptive Statistics by Experimental Arm

	(1) Control		(2) Free		(3) MF		(4) All		p-value*
A: Households characteristics									
BISWA Network[a]	0.667	(0.318)	0.630	(0.335)	0.629	(0.300)	0.642	(0.318)	0.624
Close BISWA Network[b]	0.459	(0.343)	0.442	(0.335)	0.405	(0.305)	0.435	(0.328)	0.442
Influential BISWA Network[c]	0.499	(0.373)	0.467	(0.370)	0.403	(0.340)	0.455	(0.363)	0.107
Household size	5.192	(2.426)	5.342	(2.570)	5.203	(2.252)	5.247	(2.417)	0.709
Log monthly expenditure per capita	6.549	(0.641)	6.520	(0.667)	6.550	(0.695)	6.540	(0.668)	0.869
Fraction of males	0.511	(0.174)	0.485	(0.178)	0.508	(0.170)	0.501	(0.174)	0.915
# Children under 5	0.487	(0.787)	0.537	(0.806)	0.534	(0.798)	0.520	(0.797)	0.664
# Members, completed some schooling	3.468	(2.275)	3.485	(2.326)	3.453	(2.130)	3.469	(2.224)	0.985
# Rooms in the dwelling	2.995	(1.721)	3.169	(2.060)	3.059	(1.900)	3.075	(1.902)	0.647
Access to electricity	0.402	(0.491)	0.417	(0.494)	0.402	(0.491)	0.407	(0.492)	0.953
Scheduled tribe/caste	0.745	(0.745)	0.790	(0.790)	0.725	(0.725)	0.754	(0.431)	0.141
Observations	369		391		393		1153		
B: Village Characteristics									
Number of households	261.4	(327.7)	359	(501.6)	284.3	(0.484)	301.5	(383.8)	0.534
Total population (persons)	1180.6	(1483.2)	1664.3	(2416.5)	1258.4	(1312.9)	1367.8	(1803.7)	0.496
% BISWA members	22.51	(21.97)	20.53	(20.23)	19.98	(16.13)	21.00	(19.48)	0.894
Area of Village (in hectares)	413.1	(343.7)	476.4	(340.9)	417.4	(388.25)	435.64	(356.903)	0.615
Medical Facilities (Available/NA)	0.255	(0.441)	0.298	(0.462)	0.255	(0.441)	0.270	(0.445)	0.872
Forests: % village area	0.125	(0.174)	0.087	(0.125)	0.082	(0.143)	0.097	(0.149)	0.367
Irrigated area: % village area	0.151	(0.244)	0.188	(0.265)	0.183	(0.268)	0.174	(0.258)	0.733
Non-irrigated area: % village area	0.504	(0.242)	0.483	(0.250)	0.510	(0.283)	0.499	(0.257)	0.867
Observations	47		47		47		141		

Panel A: Data from winter 2008-09 from a supplemental sample of 1,153 households non affiliated to BISWA. Panel B: Data on village characteristics from Census of India, 2001. Standard deviations in parentheses. * p-value for the joint test of equality across the three experimental arms (robust to intra-village correlation of residuals. [a] BISWA Network is calculated as the fraction of baseline households from the same village known to the respondent's households. [b] Close BISWA Network is defined as the fraction of baseline households with whom the respondent's households interacts at least once a week. [c] Influential BISWA Network is defined as the fraction of baseline households from the same village whose opinions about ways of protecting oneself from malaria are taken into account by the respondent's household.

Table 2.3: Outcome Variables, by Experimental Arm

	(1) Control		(2) Free		(3) MF		(4) All		(5) p-value
Recently acquired at least one ITN[1]	0.092	(0.290)	0.072	(0.258)	0.142	(0.350)	0.102	(0.303)	0.089*
Recently acquired at least one bednet[2]	0.371	(0.484)	0.373	(0.484)	0.427	(0.495)	0.391	(0.488)	0.401
# recently acquired ITNs	0.192	(0.666)	0.123	(0.491)	0.298	(0.866)	0.205	(0.696)	0.079*
# recently acquired bednets	0.740	(1.164)	0.770	(1.312)	0.913	(1.314)	0.809	(1.269)	0.298
# recently purchased bednets	0.596	(1.072)	0.619	(1.243)	0.710	(1.222)	0.643	(1.183)	0.431
Total number of ITNs owned	0.352	(1.091)	0.302	(0.937)	0.461	(1.138)	0.372	(1.060)	0.382
Total number of bednets owned	1.621	(1.782)	1.601	(1.858)	1.784	(1.753)	1.670	(1.799)	0.743
% of members used ITN last night	0.039	(0.180)	0.043	(0.187)	0.047	(0.185)	0.043	(0.184)	0.382
% of members used any net last night	0.060	(0.269)	0.070	(0.279)	0.084	(0.334)	0.072	(0.296)	0.418
Malaria prevalence[3]	0.197	(0.398)	0.208	(0.406)	0.206	(0.404)	0.203	(0.403)	0.957
Hemoglobin[4]	11.5	(1.96)	11.4	(1.98)	11.4	(1.93)	11.5	(1.95)	0.628
Observations	369		391		393		1153		

Data from winter 2008-09 from a supplemental sample of 1,153 households non affiliated to BISWA. Standard deviations are reported in parentheses. The p-value in column (5) are for the joint test of equality across the three experimental arms (robust to intra-village correlation of residuals). * Indicates rejection at the 10% significance level. Notes:

[1] = 1 if the household purchased at least one ITN/bednet in the past 18 months.

[2] "Bednets" includes both treated and untreated bed nets.

[3] Malaria prevalence was measured using blood tests from 2,345 individuals of age below 10 or 18-40 from 876 households.

[4] Hemoglobin was measured using blood tests from 2,362 individuals of age below 10 or 18-40 from 881 households.

Table 2.4: Cross-arm Differences in Outcomes as a Function of Links to BISWA Households

Dependent variables	(1) Recently acquired at least one ITN	(2) Fraction of household members slept under ITN last night	(3) Recently acquired at least one net	(4) Fraction of household members slept under net last night
BISWA Network $\hat{\alpha}$	0.051	-0.043	0.050	-0.044
	(0.052)	(0.039)	(0.091)	(0.042)
BISWA Network×Free $\hat{\beta}$	-0.008	0.071*	0.183	0.157*
	(0.068)	(0.042)	(0.140)	(0.087)
Observations	759	759	759	759
R-squared	0.289	0.287	0.246	0.315

Data from winter 2008-09 from a supplemental sample of households non affiliated to BISWA. Only households in Free and Control village are included. Robust standard errors clustered at village level are reported in parentheses. All specifications include village fixed effects and the following control variables: household size, number of children under 5 years old, fraction of males in the household, number of household members who completed some schooling, number of rooms in the dwelling, access to electricity, and whether the household belonged to a scheduled tribe/caste. In columns 2 and 4, household-level observations on bednet usage rates are weighted by household size. *** $p < 0.01$, **$p < 0.05$, *$p < 0.1$.

Table 2.5: Alternative Characterizations of BISWA Networks

Dependent variable:	Recently acquired at least one ITN		Slept under ITNs last night		Recently acquired at least one net		Slept under nets last night	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Measure of network:	Close BISWA Network	Influential BISWA Network	Close BISWA Network	Influential BISWA Network	Close BISWA Network	Influential BISWA Network	Close BISWA Network	Influential BISWA Network
BISWA Network $\hat{\alpha}$	0.031	0.001	0.001	-0.003	-0.042	0.009	-0.033	-0.019
	(0.042)	(0.039)	(0.028)	(0.031)	(0.084)	(0.078)	(0.059)	(0.044)
BISWA Network × Free $\hat{\beta}$	-0.007	0.054	0.019	-0.009	0.183	0.212*	0.147	0.084
	(0.063)	(0.056)	(0.038)	(0.036)	(0.128)	(0.123)	(0.097)	(0.075)
Observations	759	759	759	759	759	759	759	759
R-squared	0.288	0.289	0.285	0.285	0.242	0.248	0.316	0.312

Data from winter 2008-09 from a supplemental sample of households non affiliated to BISWA. Only households in Free and Control village are included. Robust standard errors clustered at village level are reported in parentheses. All specifications include village fixed effects and the following control variables: household size, number of children under 5 years old, fraction of males in the household, number of household members who completed some schooling, number of rooms in the dwelling, access to electricity, and whether the household belonged to a scheduled tribe/caste. Close BISWA Network is defined as the fraction of baseline households with whom the respondent's households interacts at least once a week. Influential BISWA Network is defined as the fraction of baseline households from the same village whose opinions about ways of protecting oneself from malaria are taken into account by the respondent's household.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Table 2.6: Peer Effects in Bednet Ownership and Usage

Dependent variable:	Recently purchased at least one ITN		Slept under ITNs last night		Recently purchased at least one net		Slept under nets last night	
	OLS	IV	OLS	IV	OLS	IV	OLS	IV
A:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average per capita bednets	0.151**	-0.050	0.063**	0.020	0.346***	0.042	0.102*	0.056
owned by peers	(0.071)	(0.138)	(0.030)	(0.084)	(0.077)	(0.200)	(0.054)	(0.139)
Constant	0.032	0.126*	0.014	0.034	0.230***	0.372***	0.024	0.045
	(0.031)	(0.069)	(0.014)	(0.039)	(0.041)	(0.098)	(0.027)	(0.066)
First stage F stat		31.55		31.55		31.55		31.55
Hansen J Stat		3.817		0.100		1.848		0.416
p-value		(0.051)		(0.751)		(0.174)		(0.519)
Observations	1153	1153	1153	1153	1153	1153	1153	1153
B:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Average last night ITN usage	0.007	-0.107	0.024	0.007	0.071	-0.036	0.046	0.016
among peers	(0.046)	(0.090)	(0.023)	(0.062)	(0.069)	(0.137)	(0.043)	(0.100)
Constant	0.100***	0.134***	0.036***	0.041**	0.370***	0.402***	0.058***	0.067**
	(0.021)	(0.034)	(0.009)	(0.019)	(0.032)	(0.048)	(0.018)	(0.032)
First stage F stat		40.47		40.47		40.47		40.47
Hansen J Stat		2.809		0.154		1.634		0.558
p-value		(0.094)		(0.694)		(0.201)		(0.455)
Observations	1153	1153	1153	1153	1153	1153	1153	1153

Data from winter 2008-09 from a supplemental sample of households non affiliated to BISWA. Robust standard errors clustered at village level are reported in parentheses. *** $p< 0.01$, ** $p < 0.05$, * $p < 0.1$.

Bibliography

Acemoglu, D., Johnson, S., and Robinson, J. (2001), "The Colonial Origins of Comparative Development: An Empirical Investigation," *American Economic Review*, 91, 1369–1401.

Alatas, V., Banerjee, A., Chandrasekhar, A., Hanna, R., and Olken, B. (2012), "Network Structure and the aggregation of information theory and evidence from Indonesia," Working Paper.

Alesina, A., Baqir, R., and Easterly, W. (1999), "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 114, 1243–1284.

Altonji, J., Hayashi, F., , and Kotlikoff, L. (1992), "Is the Extended Family Altruistically Linked? Direct Tests Using Micro Data," .

Anderson, L., Mellor, J., and Milyo, J. (2004), "Social Capital and Contributions in a Public-Goods Experiment," *American Economic Review*, 94, 373–376.

Angelucci, M. and Giorgi, G. D. (2009), "Indirect Effects of an Aid Program: How Do Cash Transfers Affect Ineligibles' Consumption?" *American Economic Review*, 99, 486–508.

Arcand, J.-L. and Jaimovich, D. (2012), "Does ethnic diversity decrease economic interactions? Evidence from exchange networks in rural Gambia," Working Paper.

Arcand, J.-L., Cen, Y.-P., He, Y., Diop, C. I. F., Wouabe, E. D., Garbouj, M., Jaimovich, D., and Zec, S. (2010), "The Gambia CDDP baseline: rural household survey, qualitative survey, village network survey," Tech. rep., The Graduate Institute, Geneva, Switzerland.

Ballester, C., Calv-Armengol, A., and Zenou, Y. (2006), "Who's Who in Networks. Wanted: The Key Player," *Econometrica*, 74, 1403–1417.

Bandiera, O. and Rasul, I. (2006), "Social Networks and Technology Adoption in Northern Mozambique," *The Economic Journal*, 116, 869–902.

Banerjee, A., Chandrasekhar, A., Duflo, E., and Jackson, M. (2013), "The Diffusion of Microfinance," .

Belhaj, M., Bramoullé, Y., and Deroian, F. (2012), "Network Games under Strategic Complementarities," Working Paper.

Bertrand, M., Luttmer, E., and Mullainathan, S. (2000), "Network Effects and Welfare Cultures," *Quarterly Journal of Economics*, 115, 1019–1055.

Besley, T. and Case, A. (1994), "Diffusion as a learning process: Evidence from HYV cotton," Princeton University, RPDS Discussion Paper 174.

Besley, T. and Coate, S. (1995), "Group lending, repayment incentives and social collateral," *Journal of Development Economics*, 46, 1–18.

Bloch, F. and Zenginobuz, U. (2007), "The effect of spillovers on the provision of local public goods," *Review of Economic Design*, 11, 199–216.

Bonacich, P. (1972), "Factoring and weighting approaches to clique identification," *Journal of Mathematical Sociology*, 2, 113–120.

Bonacich, P. (1987), "Power and Centrality: A Family of Measures," *American Journal of Sociology*, 92, 1170–1182.

Bonacich, P. (2007), "Some unique properties of eigenvector centrality," *Social Networks*, 29, 555–564.

Bramoullé, Y. and Kranton, R. (2007), "Public goods in Networks," *Journal of Economic Theory*, 135, 478–494.

Bramoullé, Y., Djebbari, H., and Fortin, B. (2009), "Identification of peer effects through social networks," *Journal of Econometrics*, 150, 41–55.

Bramoullé, Y., Kranton, R., and D'Amours, M. (2013), "Strategic Interactions and Networks," *American Economic Review*, Forthcoming.

Brock, W. and Durlauf, S. (2001), *Interactions-Based Models*, vol. 5 of *Handbook of Econometrics*, chap. 54, Elsevier Science BV.

Carpenter, J., Kariv, S., and Schotter, A. (2010), "Network Architecture and Mutual Monitoring in Public Goods Experiments," IZA Discussion Paper No. 5307.

Conley, T. G. and Udry, C. R. (2010), "Learning About a New Technology: Pineapple in Ghana," *American Economic Review*, 100, 35–69.

Cox, D. and Fafchamps, M. (2008), "Extended Family and Kinship Networks: Economic Insights and Evolutionary Directions," in *Handbook of Development Economics*, eds. P. Schultz and J. Strauss, vol. 4, chap. 58, pp. 3711–3784, Elsevier Science.

De Giorgi, G., Pellizzari, M., and Redaelli, S. (2009), "Identification of Social Interactions through Partially Overlapping Peer Groups," *American Economic Journal: Applied Economics*, 2, 241–75.

Dupas, P. (2010), "Short-Run Subsidies and Long-Run Adoption of New Health Products: Evidence from a Field Experiment," Working Paper.

Dupas, P. (2012), "Health Behavior in Developing Countries," *Annual Review of Economics*, 3, 425–449.

Durlauf, S. and Fafchamps, M. (2005), "Social Capital," in *Handbook of Economic Growth*, eds. P. Aghion and S. Durlauf, vol. IV, chap. 36, pp. 1639–1699, Elsevier Science.

Easterly, W. and Levine, R. (1997), "Africas growth tragedy: policies and ethnic divisions," *QJE*, 112, 12031250.

Ertana, A., Pageb, T., and Puttermanc, L. (2009), "Who to punish? Individual decisions and majority rule in mitigating the free rider problem," *European Economic Review*, 53, 495511.

Esteban, J., Mayoral, L., and Ray, D. (2012), "Ethnicity and Conflict: An Empirical Study," *American Economic Review*, 102, 1301–1342.

Fearon, J. (2003), "Ethnic and Cultural Diversity by Country," *Journal of Economic Growth*, 8, 195222.

Fehr, E. and Gaechter, S. (2000), "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, 90, 980–994.

Foster, A. D. and Rosenzweig, M. R. (1995), "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 103, 1176–1209.

Galeotti, A. and Goyal, S. (2010), "The Law of the Few," *American Economic Review*, 100, 1468–92.

Gamble, D. (1949), *Gambia Studies Series*, Research Department, Colonial Office, London.

Graham, B. (2008), "Identifying social interactions through conditional variance restrictions," *Econometrica*, 76, 643–660.

Habyarimana, J., Humphreys, M., Posner, D., and Weinstein, J. (2007), "Why Does Ethnic Diversity Undermine Public Goods Provision?" *American Political Science Review*, 101, 709–725.

Hawley, W. A., Phillips-Howard, P. A., ter Kuile, F. O., Terlouw, D. J., Vulule, J. M., Ombok, M., Nahlen, B. L., Gimnig, J. E., Kariuki, S. K., Kolczak, M. S., and Hightower, A. W. (2003), "Community-wide effects of permethrin-treated bed nets on child mortality and malaria morbidity in western Kenya." *The American Journal of Tropical Medicine and Hygiene*, 68, 121–7.

Holla, A. and Kremer, M. (2009), "Pricing and Access: Lessons from Randomized Evaluations in Education and Health," Center for Global Development Working Paper 158.

Jackson, M. (2007), *Social and Economic Networks*, Princeton University Press, Princeton, New Jersey.

Jackson, M., Rodriguez-Barraquer, T., and Tan, X. (2012), "Social Capital and Social Quilts: Network Patterns of Favor Exchange," *American Economic Review, Forthcoming*.

Jaimovich, D. (2011), "Macrostructure and microstructure: Evidence from overlapping village networks in the Gambia," MPRA Working Paper No. 38932.

Jaimovich, D. (2013), "Missing links, missing markets: Evidence from exchange networks in rural Gambia," MPRA Working Paper No. 44080.

Jarju, P. O. (2009), "National report on Adaptation of water resources in the Gambia," Department of Water Resources, Government of the Gambia.

Killeen, G. F., Smith, T. A., Ferguson, H. M., Mshinda, H., Abdulla, S., Lengeler, C., and Kachur, S. P. (2007), "Preventing childhood malaria in Africa by protecting adults from mosquitoes with insecticide-treated nets," *PLoS Medicine*, 4.

Kremer, M. and Miguel, E. (2007), "The illusion of sustainability," *Quarterly Journal of Economics*, 122, 1007–1065.

Kremer, M., , Miguel, E., Mullainathan, S., Null, C., and Zwane, A. P. (2011), "Social Engineering: Evidence from a suite of take-up experiments in Kenya," Working Paper.

Kumar, A., Valecha, N., Jain, T., and Dash, A. P. (2007), "Burden of Malaria in India: Retrospective and Prospective View," *American Journal of Tropical Medical Hygiene*, 77, 69–78.

Lengeler, C. (2004), "Insecticide-treated bed nets and curtains for preventing malaria." *Cochrane database of systematic reviews (Online)*, Issue 2. Art. No.: CD000363. DOI: 10.1002/14651858.CD000363.pub2.

Manski, C. F. (1993a), "Identification of endogenous social effects: The reflection problem," *Review of Economic Studies*, 60, 531–542.

Manski, C. F. (1993b), "Identification of endogenous social effects: The reflection problem," *Review of Economic Studies*, 60, 531–542.

Masclet, D., Noussair, C., Tucker, S., and Villeval, M.-C. (2003), "Monetary and Nonmonetary Punishment in the Voluntary Contributions Mechanism," *American Economic Review*, 93, 366–380.

Miguel, E. and Kremer, M. (2004), "Worms: identifying impacts on education and health the presence of treatment externalities," *Econometrica*, 72, 159–217.

Miguel, M. and Gugerty, M. K. (2005), "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya," *Journal of Public Economics*, 89, 2325–68.

Munshi, K. and Myaux, J. (2006), "Social norms and the fertility transition," *Journal of Development Economics*, 80, 1–38.

Mwakikagile, G. (2010), *The Gambia and Its People: Ethnic Identities and Cultural Integration in Africa*, New Africa Press, Tanzania.

Nowell, C. and Tinkler, S. (1994), "The influence of gender on the provision of a public good," *Journal of Economic Behavior and Organization*, 25, 25–36.

Sonko-Godwin, P. (2003), *Ethnic groups of the Senegambia region: A brief history*, Sunrise Publihers, Kanifing, the Gambia, 3rd edn.

Stock, J. and Yogo, M. (2005), *Testing for Weak Instruments in Linear IV Regression*, Cambridge University Press, New York.

Sylla, F. (2010), "Revitalization of the Groundnut Sector in West Africa," Global agricultural information network report, USDA Foreign Agricultural Service, Dakar, Senegal.

Tarozzi, A., Mahajan, A., Blackburn, B., Kopf, D., Krishnan, L., and Yoong, J. (2011), "Micro-loans, bednets and malaria: Evidence from a randomized controlled trial in Orissa (India)," Working Paper.

Thomas, D., Frankenberg, E., Friedman, J., Habicht, J.-P., Hakimi, M., Ingwersen, N., Jaswadi, Jones, N., McKelvey, K., Pelto, G., Sikoki, B., Seeman, T., Smith, J. P., Sumantri, C., Suriastini, W., and Wilopo, S. (2006), "Causal effect of health on labor market outcomes: Experimental evidence," Working Paper.

Valente, T., Coronges, K., Lakon, C., and Costenbader, E. (2008), "How Correlated Are Network Centrality Measures?" *Connect (Tor.)*, 28, 16–26.

van de Walle, E. and Gaye, A. (2006), "Household structure, polygyny and ethnicity in Senegambia : a comparison of census methodologies," in *African households : censuses and surveys*, ed. E. van de Walle, chap. 1, pp. 3–21, M.E. Sharpe, Inc., New York.

Wolitzky, A. (2013), "Cooperation with Network Monitoring," *Review of Economic Studies*, 80, 395–427.

Wong, S. (2012), "What have been the impacts of World Bank Community-Driven Development Programs? CDD impact evaluation review and operational and research implications." Report, The Worldbank, Washington, D.C.

World Bank (2008), "Global Purchasing Power Parities and Real Expenditures, 2005," Tech. rep., International Comparison Program, Washington DC: World Bank, Available at `www.worldbank.org/data/icp`.

Biography

Chutima "Gift" Tontarawongsa was born in Bangkok, Thailand, on 10 April 1987. She spent her childhood in her hometown, Chachoengsao, until the age of 15. She then attended Mahidol Wittayanisorn School, a boarding school in Nakorn Pathom. Upon completion of high school, Gift came to the US to attend Lafayette College in Easton, Pennsylvania where she graduated in May 2009 *summa cum laude* with a Bachelor of Arts in Economics and Mathematics. Gift began her graduate studies at Duke in the Fall of 2009, following her graduation.